

# A LAPLACE PRINCIPLE FOR HERMITIAN BROWNIAN MOTION AND FREE ENTROPY

YOANN DABROWSKI

*Dedicated to Professor Dan Virgil Voiculescu.*

**ABSTRACT.** We prove that the  $\limsup$  and  $\liminf$  variants of Voiculescu's free entropy coincide. This is based on a Laplace principle (implying a full large deviation principle) for hermitian brownian motion on  $[0, 1]$ . As a consequence, we show that microstates free entropy  $\chi(X_1, \dots, X_m)$  and non-microstate free entropy  $\chi^*(X_1, \dots, X_m)$  coincide for self-adjoint variables  $(X_1, \dots, X_m)$  satisfying a Schwinger-Dyson equation for subquadratic, bounded below, strictly convex potentials with Lipschitz derivative sufficiently approximable by non-commutative polynomials. Applying the contraction principle, we obtain a large deviation result for Haar unitaries and deduce the most general additivity property for a new extended definition of orbital free entropy. Our results are based on Dupuis-Ellis weak convergence approach to large deviations where one shows a Laplace principle in obtaining a stochastic control formulation for exponential functionals. In the non-commutative context, ultrapower analysis replaces weak-convergence of the stochastic control problems.

## 1. INTRODUCTION

In a fundamental series of papers [V2, V3, V4, V5, V6], Voiculescu introduced analogs of entropy and Fisher information in the context of free probability theory. A first microstates free entropy  $\chi(X_1, \dots, X_m)$  is defined as a normalized limit of the volume of sets of microstates i.e. matricial approximants (in moments) of the  $n$ -tuple of self-adjoints  $X_i$  living in a (tracial)  $W^*$ -probability space  $M$ . Starting from a definition of a free Fisher information [V5], Voiculescu also defined a non-microstate free entropy  $\chi^*(X_1, \dots, X_m)$ , known by the fundamental work [BCG] not to be smaller than the previous microstates entropy, and believed to be equal (at least modulo Connes' embedding conjecture). For more details, we refer the reader to the survey [V] for a list of properties as well as applications of free entropies in the theory of von Neumann algebras. The technical definitions are recalled later in subsection 2.8.

As pointed out in the review article [V], the study of free entropy has been faced with several technical questions, among which the two most famous are the equality of microstates and non-microstate definitions (the so-called unification problem [V, p22]) and the equality of two variants of free entropy with a  $\limsup$ , we will call  $\chi$ , or a  $\liminf$  over the size of matrix approximations  $N$  [V, Rmk (a) p9]. We will call  $\underline{\chi}$  the  $\liminf$  variant. In absence of equality, several natural questions, most notably additivity of free entropy for free variables and the free entropy power inequality [V1], were solved using yet another ultrafilter variant  $\chi^\omega$  in between the previous two definitions. Our goal in this paper is to solve completely this second question about limits and partially the unification problem.

Moreover, while non-microstate free entropy came naturally, from its very start in [V5], with a notion of relative entropy  $\chi^*(X_1, \dots, X_m : B)$ , for variables relative to a von Neumann algebra  $B$ , various competing definitions remained in the microstates picture from the original definition in [V2] (for  $B = W^*(Y_1, \dots, Y_n)$ ), the technically useful entropy in presence [V3] and a final attempt in [S02]. At first, our results point out this last definition  $\chi(X_1, \dots, X_m | B)$  also recalled in section 2.8 as the closest one to  $\chi^*(X_1, \dots, X_n : B)$ , but working more with orbital entropy, we finally show this definition also coincides with the first definition in [V2], (see (1.6) in Theorem E).

Let us state our main result about equality of microstates variants in this context:

---

2010 *Mathematics Subject Classification.* Primary 46L54, 60F10; Secondary 60B20, 60G15, 46M07.

**Theorem A.** *Let  $X_1, \dots, X_m, m \geq 1$  self-adjoint variables in a finite von Neumann algebra  $(M, \tau)$ . Let  $B \subset M$  an  $R^\omega$ -embeddable von Neumann sub-algebra. Then we have equality of the  $\liminf$  and  $\limsup$  variants of relative free entropy:*

$$\chi(X_1, \dots, X_m | B) = \underline{\chi}(X_1, \dots, X_m | B).$$

The proof will be given at the end of subsection 7.1.

As an immediate application of [V1] and [BD13], we obtain the most satisfactory formulation of the relation of free entropy and freeness and of the free entropy power inequality, without the technical use of ultrafilters.

**Corollary B.** *If  $\chi(X_1, \dots, X_n, Y_1, \dots, Y_m) > -\infty$ , then*

$$\chi(X_1, \dots, X_n, Y_1, \dots, Y_m) = \chi(X_1, \dots, X_n) + \chi(Y_1, \dots, Y_m)$$

*if and only if  $\{X_1, \dots, X_n\}$  and  $\{Y_1, \dots, Y_m\}$  are free. In this case, if  $n = m$ , we have:*

$$\exp\left(\frac{2}{n}\chi(X_1 + Y_1, \dots, X_n + Y_n)\right) \geq \exp\left(\frac{2}{n}\chi(X_1, \dots, X_n)\right) + \exp\left(\frac{2}{n}\chi(Y_1, \dots, Y_n)\right).$$

*Proof.* The only if part is our result in [BD13, Corol 7.4]. The ultrafilter variant of the if part and the entropy-power inequality is from [V1, Th 3.8 and 3.9]. The equality  $\chi^\omega = \chi$  deduced from our Theorem A concludes.  $\square$

Our approach to prove Theorem A is based on large deviations for hermitian brownian motion as in [CDG1, CDG2, BCG, GZ]. In the several variable case, the best result in [BCG] gave a large deviation upper bound and a large deviation lower bound with two (a priori) different rate functions. Our main improvement is to be able to get the same rate function, which is not unfortunately always known to be related to non-microstate free entropy  $\chi^*$ . Having the same (good) rate function is the key to the equality of  $\liminf$  and  $\limsup$  variants and looking for  $\chi^*$ , or a too similar formula, was one of the main obstacles in solving this question in our viewpoint. For the terminology and methodology of large deviation Theory we refer to [DE]. Since we follow their weak-convergence approach (that should maybe be called stochastic control approach), we also follow their terminology ‘‘Laplace principle’’. We will refer to [DZ] for more technical results.

**Theorem C.** *Fix a sequence  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$  of (deterministic) unitary matrices. Assume that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\nu$ , law of unitary variables  $\nu$ . We also assume either  $m \geq 2$ , or  $m = 1$  and the von Neumann algebra  $W^*(\nu)$  generated by  $\nu$  is diffuse.*

*Let  $H_t^{N,1}, \dots, H_t^{N,m}$   $t \in [0, 1]$  be  $m$  independent  $N \times N$  hermitian brownian motions with usual GUE normalisation  $E(\frac{1}{N}\text{Tr}((H_t^{N,j})^2)) = t$ . Let  $\hat{\sigma}_{\Upsilon_N}^N$  the joint law of the unitary processes  $u(H_t^{N,j}) = \frac{H_t^{N,j} + 4i}{H_t^{N,j} - 4i}$  and the unitaries  $\Upsilon_N$  in  $(\mathcal{T}(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\nu^d), d = d_{0,0})$ , the tracial state space for the universal  $C^*$ -algebra with  $m$  unitary processes and  $\mu_\nu$  unitaries (see subsection 2.1 for the distance we put on it).*

*Then,  $\hat{\sigma}_{\Upsilon_N}^N$  satisfies a Laplace principle in the scale  $N^{-2}$  in  $(\mathcal{T}(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d)$  (and thus a Large deviation principle) with the good rate function  $I_\nu : (\mathcal{T}(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d) \rightarrow [0, \infty]$ , depending only on the law  $\mu_\nu$  of  $\nu$  and not of the sequence  $(\Upsilon_N)$  and given by  $I_\nu(\tau) = \sup_{f \in C_b^0} -f(\tau) + \Lambda(\nu, f)$ . Moreover, on the space of continuous bounded functions involved  $C_b^0 = C_b^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , we have the formula:*

$$\lim_{N \rightarrow \infty} -\frac{1}{N^2} \log E\left(e^{-N^2 f(\hat{\sigma}_{\Upsilon_N}^N)}\right) = \Lambda(\nu, f) := \inf \left\{ \frac{1}{2} \int_0^1 \|u_s\|_2^2 ds + f(\tau_{S + \int_0^\cdot u_s ds, \nu}) : u \in \mathcal{P}_\nu \right\}$$

*where  $S = (S^1, \dots, S^m)$  are  $m$  standard free brownian motions free from  $\nu$  in  $(M, \tau)$ , and  $\mathcal{P}_\nu$  is the set of control processes in  $L^2([0, 1], L^2(M))$  adapted to  $\mathcal{F}_s = W^*(\nu, S_u, u \leq s)$ .*

This large deviation result is obtained in a more technical form in Theorem 6.2 and Corollary 6.3. Note that the assumption  $W^*(\nu)$  diffuse excludes the result in [BAG] when  $m = 1$ . The key

part in its proof is to work harder on the Large Deviation upper bound in order to get the rate function as a minimization problem on a smaller set in order to get more easily a lower bound with the same rate function. It does not use any approximation of variables as in the final result in the one variable case [GZ2]. Apart from the exponential tightness and the identification of the brownian bridge as reaching an infimum, it does not use much of the argument in [BCG]. We obtained above a variational formula for the rate function in the spirit of a free analogue of pressure defined recently by Hiai [Hi], but in a non-convex setting with the use of Bryc's inverse Varadhan lemma of large deviation theory (and this is necessary since Voiculescu's entropy is not concave and we don't use its concavification [BD13]). The formula is based on a free analogue of Boué-Dupuis formula [BD] for free pressure and is based on a recent improvement by Üstünel [Us14] of the original formula, better suited for convex analysis, and applied to hermitian brownian motion (see also [L] for a nice introduction and other applications of this formula). This is our starting point for our use of the weak-convergence approach of Dupuis and Ellis (that motivated the discovery of their formula in [BD]) that shows a Laplace principle (equivalent to large deviation principle, LDP for short, under appropriate assumptions) based on limits of a stochastic control problem. In our version, the weak-convergence part is replaced by ultraproduct analysis (see Subsection 2.9 for a short introduction and references).

Technically, our result is strongly based on convex analysis, this is why we can still get a partial answer about the unification problem under a convexity assumption. Note that beyond the general inequality in [BCG], and the free product of the single variable case, not much was known about the equality  $\chi = \chi^*$ . Even when an explicit computation of microstates free entropy was known as for small perturbation of semicircular variables [GM], or when implicit formulas are known thanks to the recent developments of free transport [GS14], no equality cases was known even for perturbations of semi-circular variables, to the best of our knowledge. This is due to the technicality of the definition of non-microstate free entropy in [V5] for which no non-linear change of variable is known. Much more cases of equality were known for the related microstates free entropy dimension [MS, D10]. We have a similar result relative to  $B$  with a technical assumption we will state as Theorem 7.4 and which is similar to the one in Theorem C. Since the notion of Schwinger-Dyson equation is less known in this case, and only explained in Theorem 4.4, we stick to the case similar to [GS09] (with a much less constraining notion of convexity defined in subsection 2.2). The exact formulation of the function space for the potential  $\mathcal{E}_{app}^{1,1}(\mathcal{T}_2(\mathcal{F}_1^m), d_2)$  will be given in this subsection. A basic example of function, linear in the trace, is for instance for  $\lambda$  small enough to get convexity:

$$g(\tau) = D + C \sum_{j=1}^k \sum_{l=1}^m \tau((4i \frac{u_j^l + 1}{u_j^l - 1})^* (4i \frac{u_j^l + 1}{u_j^l - 1})) + \Re(\lambda \tau((u_{j_1}^{l_1})^{\epsilon_1} \dots (u_{j_m}^{l_m})^{\epsilon_m}))$$

for the trace expressed in unitary variables  $u_j^l = u(X_{t_j}^i)$   $t_j \in [0, 1]$ ,  $\epsilon_i \in \{+1, -1\}$

In words, this is a space of convex bounded bellow subquadratic functions with lipschitz derivative and a natural approximation of the derivative by non-commutative polynomials. This is because we give enough explicit examples in this space in this subsection that we can prove Theorem A. Finally, we also deduce various other conjectures for free entropy for semi-circular perturbation in this case, most notably the technical equality of entropy to entropy in presence  $\chi(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m : \sqrt{t}S_1, \dots, \sqrt{t}S_m)$ , appearing in the definition of microstates free entropy dimension, and the continuity of free Fisher information  $\Phi^*$  under semicircular perturbation (see the question in [V, p23]).

**Theorem D.** *Let  $m \geq 2$  and  $V \in \mathcal{E}_{app}^{1,1}(\mathcal{T}_2(\mathcal{F}_1^m), d_2)$ . Let  $X_1, \dots, X_m$  having law  $\tau_V$ , the unique solution of  $(SD_V)$  obtained in Theorem 4.4 (or equivalently a conjugate variable in the sense of [V5] given by usual cyclic derivatives  $X_i + \mathcal{D}_i V$ ). Then, we have the equality:*

$$\chi(X_1, \dots, X_m) = \chi^*(X_1, \dots, X_m).$$

*If moreover  $S_1, \dots, S_m$  is a free semicircular system free from  $\{X_1, \dots, X_m\}$ , then for any  $t > 0$ :*

$$\chi(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m : \sqrt{t}S_1, \dots, \sqrt{t}S_m) = \chi^*(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m),$$

the free Fisher information  $t \mapsto \Phi^*(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m)$  is Hölder continuous (of exponent  $1/2$ ) on  $[0, \infty)$  and for  $(\xi_{1,t}, \dots, \xi_{m,t})$  the conjugate variables in  $W^*(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m)$  and  $\delta_t$  the corresponding free difference quotient, we have the integral formula for  $0 \leq s < t$ :

$$\Phi^*(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m) = \Phi^*(X_1 + \sqrt{s}S_1, \dots, X_m + \sqrt{s}S_m) - \int_s^t \sum_{i=1}^m \|\delta_u(\xi_{i,u})\|_2^2 du.$$

The last equation is an improvement of an inequality in [D14]. The equality version enables to compute the derivative of free Fisher information along free brownian motion.

Let us point out that contrary to the advertised potential application of our kind of full large deviation principle in [V, section 3.8], our large deviation result does not imply Connes' embedding conjecture. This would require having equality  $\chi = \chi^*$  for more general variables, which is naturally a desirable research project. The previous result suggests that the various conjectures in [V] that we solve for variables coming from convex potentials are strongly related to the unification problem and maybe to Connes' embedding conjecture.

In the final paper [V6] of his series introducing free entropy, Voiculescu developed a notion of free mutual information, the free analogue of mutual information (the entropy of a measure on a product space relative to the tensor product of its marginals). Voiculescu conjectures various additivity relations with free entropy (see e.g. [V]) but none is known in the non-microstate setting. Later, [HMU] introduced the microstates variant  $\chi_{orb}(\mathbf{X}_1, \dots, \mathbf{X}_m)$  for a bunch of sets of self-adjoint variables  $\mathbf{X}_i = (X_{i1}, \dots, X_{i\nu})$  when  $W^*(\mathbf{X}_i)$  is hyperfinite. They proved a first additivity result with free entropy in the one variable case  $\nu = 1$  in the form:

$$(1.1) \quad \chi(X_1, \dots, X_n) = \chi_{orb}(X_1, \dots, X_n) + \chi(X_1) + \dots + \chi(X_n).$$

They also show the important property that  $\chi_{orb}(\mathbf{X}_1, \dots, \mathbf{X}_m)$  depends only on  $W^*(\mathbf{X}_i)$  in the hyperfinite case. In [BD13] and [Ue14], several attempts were made of finding a generalization of orbital entropy to non-hyperfinite sets of variables but they failed to recover both properties: dependence on algebra and full additivity. Since orbital entropy is based on Haar measure, we can use a large deviation principle for Haar Measures to study it. We will obtain this LDP in section 8.1 (cf. Theorem 8.3) as an application of the contraction principle. Then we can extend orbital free entropy to sets of unitary variables  $\mathbf{U}_i$  (giving the same result as before in [HMU] when  $W^*(\mathbf{X}_i) = W^*(\mathbf{U}_i)$  hyperfinite, but a priori not equal to its attempted extensions in [BD13] and [Ue14]). Our new extension (see Definition 8.6) satisfies all the desired properties (cf. Proposition 8.8 for the most basic ones) especially it answers the additivity question for general non-hyperfinite variables. Said otherwise, if we define orbital entropy using the additivity relation (1.4) when all entropic quantities are finite, it satisfies the amazing von Neumann algebraic invariance that was unlikely for free entropy  $\chi$ .

**Theorem E.** *For finite sets of unitary variables  $\mathbf{U}_i, v$ ,*

$$\chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) = \chi_{orb}(W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_m) | W^*(v))$$

*depends only on  $W^*(\mathbf{U}_i), W^*(v)$  and satisfy the additivity relations:*

(1) *For any  $m_1 < \dots < m_n < m$ , we have :*

$$(1.2) \quad \begin{aligned} \chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) &= \chi_{orb}(W^*(\mathbf{U}_1, \dots, \mathbf{U}_{m_1}), \dots, W^*(\mathbf{U}_{m_n+1}, \dots, \mathbf{U}_m) | W^*(v)) \\ &+ \chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_{m_1}) + \dots + \chi_{orb}(\mathbf{U}_{m_n+1}, \dots, \mathbf{U}_m). \end{aligned}$$

(2) *Especially, if  $\chi_{orb}$  bellow are finite,  $W^*(\mathbf{U}_1, \dots, \mathbf{U}_{m_1}), \dots, W^*(\mathbf{U}_{m_n+1}, \dots, \mathbf{U}_m)$  are free if and only if*

$$(1.3) \quad \begin{aligned} \chi_{orb}(W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_m)) &= \\ \chi_{orb}(W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_{m_1})) &+ \dots + \chi_{orb}(W^*(\mathbf{U}_{m_n+1}), \dots, W^*(\mathbf{U}_m)) \end{aligned}$$

*and if and only if*

$$\chi_{orb}(W^*(\mathbf{U}_1, \dots, \mathbf{U}_{m_1}), \dots, W^*(\mathbf{U}_{m_n+1}, \dots, \mathbf{U}_m)) = 0.$$

(3) For any set of self adjoint variables  $\mathbf{X}_i$ , we have :

$$(1.4) \quad \chi(\mathbf{X}_1, \dots, \mathbf{X}_m | v) = \chi_{\text{orb}}(W^*(\mathbf{X}_1), \dots, W^*(\mathbf{X}_m) | W^*(v)) + \chi(\mathbf{X}_1) + \dots + \chi(\mathbf{X}_m).$$

$$(1.5) \quad \chi(\mathbf{X}_1, \dots, \mathbf{X}_m | v) = \chi_{\text{orb}}(W^*(\mathbf{X}_1, \dots, \mathbf{X}_m), W^*(v)) + \chi(\mathbf{X}_1, \dots, \mathbf{X}_m).$$

Moreover, if  $\chi(\mathbf{X}_1, \mathbf{X}_2) > -\infty$ , then:

$$(1.6) \quad \chi(\mathbf{X}_1 | W^*(\mathbf{X}_2)) = \chi(\mathbf{X}_1, \mathbf{X}_2) - \chi(\mathbf{X}_2).$$

Let us finally give an overview of the paper. Section 2 deals with the various preliminaries and notation. The notation strongly needed in the paper about tracial state spaces is in subsection 2.1, the classes of non-commutative convex functions are defined in 2.2 and preliminaries on hermitian brownian motion in 2.4. Subsection 2.5 explains the two kinds of concentration of measure results we will use, the well-known one coming from convexity via logarithmic Sobolev inequality useful for the large deviation upper bound, and the less well-known one coming from a Poulsen Simplex property from [BD13] useful for the large deviation lower bound. The other preliminaries are less critical, subsection 2.3 deals with Malliavin calculus needed to understand the statements of [Us14], the classical expression of regularity of derivatives of convex functions via second order difference quotients (better suited in stochastic control estimates) is explained in subsection 2.6, background on classical and free entropy are gathered in 2.7 and 2.8. References on ultraproduct analysis are given in subsection 2.9.

Section 3 recalls the version of the Boué-Dupuis formula due to Üstünel that we will need. As usual in stochastic control, the value function at time  $t$  will be used to characterize the minimiser to the optimization problem and it is expressed with an optimization problem given in our case by an application of the Boué-Dupuis-Üstünel formula. We use standard ideas from optimal control to give regularity properties in time and space. If the terminal cost function  $g$  we start with is convex, the value function stay convex and if this cost function  $g$  has lipschitz first derivative, one gets upper bounds on second order difference quotients of the value function  $h_t$ . They enable to get some Hölder continuity in time for the derivative that will give us the crucial uniform regularity to use ultraproduct techniques.

Section 4 recalls solutions of Stochastic differential equations (SDE) under monotone drift assumptions we will use. It also gives the corresponding version in free probability which is new and of independent interest when using convex analysis jointly with free SDE techniques. We hope it may have future applications to free transport. The uniqueness and existence of strong solutions in the free case will be crucial to obtain a value function independent of the ultrafilter in our ultraproduct analysis and thus a real limit in our Laplace deviation with a rate function not dependent on any ultrafilter. This lack of uniqueness of potential solutions of a free SDE was the issue in [BCG] since the (a priori) different values of the lower bound rate function was mainly associated with a supplementary uniqueness assumption of such a solution not known for the upper bound. The key point of our use of convex analysis is to get this uniqueness.

Section 5 uses the solution in [Us14] of the minimization problem appearing in Boué-Dupuis formula. This is the second place where convexity is used crucially. We apply it in the hermitian brownian motion and check that our estimates are sufficiently uniform in the matrix size  $N$ . The minimizer is expressed in terms of a solution of the SDE with the drift coming from the gradient of the value function. We estimated it earlier as Lipschitz in space and Hölder continuous in time. An alternative formula is given in subsection 5.1 in terms of a solution of an SDE. This will be this formula that will insure that we control the von Neumann algebra in which lives the drift of the SDE. One of the reasons of the failure of obtaining equality of  $\chi = \chi^*$  previously was that some limit of matricial analogues of conjugate variables may still depend on entries of matrices and may not be a non-commutative functional calculus of the solution. In our framework, they could be in an ultraproduct of random matrix spaces rather than in a smaller von Neumann algebra generated by the solution. The formula in this section is the key to obtain the smaller von Neumann algebra via an existence and uniqueness of some free (forward-backward) SDE.

Section 6 then explains the Laplace principle and its equivalent large deviation version. Note that following [BD], the lower bound in our Laplace principle corresponds to the usual large deviation upper bound. Section 7 gives our applications to free entropy. Section 8 deals with the



LDP for Haar measures (as above, jointly with constant matrices converging in non-commutative distribution) and its application to orbital free entropy.

**Acknowledgments :** The author wants to thank Ivan Gentil for showing him [Ko] and discussing [Ge], which motivated the use of convex analysis to the present large deviation problem, and Alice Guionnet for suggesting to write the large deviation for Haar unitaries. The author is grateful to the organizers of the "Conference on von Neumann algebras and related topics" in January 2012 at RIMS, Kyoto University, Japan, where started this research on the applications of ultraproducts to free entropy theory started. The author also wants to thank Yoshimichi Ueda for interesting conversations on orbital free entropy, which started at that time too.

## 2. PRELIMINARIES AND NOTATION

**2.1. Tracial states with second moments.** We fix a setting similar to [BCG], except that we exploit convexity more crucially. We call  $\mathcal{F}_{[0,1]}^m$  the group universal  $C^*$ -algebra in the free group with generators  $\{u_t^i, i = 1, \dots, m, t \in [0, 1]\}$ . (note they considered only the algebra with the same notation, and we will need non algebraic functions to exploit convexity.)

We call  $\mathcal{T}(\mathcal{F}_{[0,1]}^m)$  the set of tracial states on this  $C^*$  algebra. Consider also  $\mathcal{F}_n^m$  the group universal  $C^*$ -algebra in the free group with generators  $\{u_j^i, i = 1, \dots, m, j = 2, \dots, n + 1\}$ . We start indexing at 2 to differentiate from time indices in  $[0, 1]$ .

We will also consider the universal  $C^*$ -algebra free product for instance  $\mathcal{F}_{[0,1]}^m * \mathcal{F}_n^\mu$  for  $\mu \in \mathbb{N}^*$ , of course we have for instance  $\mathcal{F}_1^m * \mathcal{F}_1^\mu \simeq \mathcal{F}_1^{m+\mu}$ . In case of  $\mathcal{F}_k^m * \mathcal{F}_\nu^\mu \subset \mathcal{F}_{k+\nu}^{m+\mu}$  we consider the  $C^*$  algebra variable with generators  $\{u_j^i, i = 1, \dots, m, j = 2, \dots, k + 1, \} \cup \{u_j^i, i = m + 1, \dots, m + \mu, j = k + 2, \dots, k + \nu + 1\}$ .

For any  $\{t_1, \dots, t_n\}$ , we have a canonical  $*$ -homomorphism  $I_{t_1, \dots, t_n} : \mathcal{F}_n^m \rightarrow \mathcal{F}_{[0,1]}^m$  with the following action on generators

$$I_{t_1, \dots, t_n}(u_j^i) = u_{t_j-1}^i.$$

We denote  $\mathcal{T}^c(\mathcal{F}_{[0,1]}^m)$  the set of continuous tracial states namely those states  $\tau$  such that for any  $n$  the map  $(t_1, \dots, t_n) \in [0, 1]^n \rightarrow \tau \circ I_{t_1, \dots, t_n}(f)$  is continuous for any  $f \in \mathcal{F}_n^m$ . Similarly we consider  $\mathcal{T}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_n^\mu)$  with continuity on the continuous times variables and using  $I_{t_1, \dots, t_k} * Id : \mathcal{F}_k^m * \mathcal{F}_n^\mu \rightarrow \mathcal{F}_{[0,1]}^m * \mathcal{F}_n^\mu$  instead of  $I_{t_1, \dots, t_n}$  in the continuity requirement.

Let  $(M, \tau)$  be a (tracial)  $W^*$ -probability space, namely a von Neumann algebra  $M$  with a faithful normal trace  $\tau$ . We will assume implicitly all our  $W^*$ -probability spaces to be tracial. We consider self-adjoint processes  $X = (X_t^i, i = 1, \dots, m, t \in [0, 1])$ , with  $X_t^i = (X_t^i)^* \in L^2(M, \tau)$ ,  $i = 1, \dots, m, t \in [0, 1]$ . We write  $\tau_X \in \mathcal{T}(\mathcal{F}_{[0,1]}^m)$  the law of

$$(u(X_t^l) := \frac{X_t^l + 4i}{X_t^l - 4i}, l = 1, \dots, m, t \in [0, 1])$$

It is not hard to see that if  $t \mapsto X_t^i$  are continuous in  $L^2(M, \tau)$ , then  $\tau_X \in \mathcal{T}^c(\mathcal{F}_{[0,1]}^m)$ . We even have  $\tau_X \in \mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m)$  the space where for any  $l$ ,

$$t \in [0, 1] \mapsto \tau((4i \frac{u_t^l + 1}{u_t^l - 1})^* (4i \frac{u_t^l + 1}{u_t^l - 1})) < \infty$$

is continuous (since  $\tau_X((4i \frac{u_t^l + 1}{u_t^l - 1})^* (4i \frac{u_t^l + 1}{u_t^l - 1})) = \tau((X_t^l)^2)$ ). We use a similar notation  $\tau_X$  for state in  $\mathcal{T}(\mathcal{F}_n^m)$ . fir  $X$  self-adjoint variables. Similarly, we call  $\mathcal{T}_2(\mathcal{F}_n^m)$  the set of tracial states with  $\tau((4i \frac{u_j^l + 1}{u_j^l - 1})^* (4i \frac{u_j^l + 1}{u_j^l - 1})) < \infty, j = 1, \dots, k$ .

We first consider on  $\mathcal{T}^c(\mathcal{F}_{[0,1]}^m)$  the same topology as in [BCG], namely the one given by the distance:

$$d(\tau_1, \tau_2) = \sum_{k=1}^{\infty} 2^{-k} \sup_{(i_1, \dots, i_m) \in [1, m]^k} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{-1, 1\}^k} \sup_{(t_1, \dots, t_k) \in [0, 1]^k} |(\tau_1 - \tau_2)((u_{t_1}^{i_1})^{\epsilon_1} \dots (u_{t_k}^{i_k})^{\epsilon_k})|.$$

We also consider on  $\mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m)$  the topology given by the distance:

$$d_2(\tau_1, \tau_2) = d(\tau_1, \tau_2) + \sup_{l=1, \dots, m} \sup_{t \in [0,1]} \left| (\tau_1 - \tau_2) \left( \left( \frac{u_t^l + 1}{u_t^l - 1} \right)^* \left( \frac{u_t^l + 1}{u_t^l - 1} \right) \right) \right|$$

$$+ \sum_{k=1}^{\infty} 2^{-k} \sup_{(l, i_1, \dots, i_m) \in \llbracket 1, m \rrbracket^{k+1}} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{-1, 1\}^k} \sup_{(t, t_1, \dots, t_k) \in [0, 1]^{k+1}} |(\tau_1 - \tau_2) \left( \left( \frac{u_t^l + 1}{u_t^l - 1} \right)^* (u_{t_1}^{i_1})^{\epsilon_1} \dots (u_{t_k}^{i_k})^{\epsilon_k} \right)|.$$

We put a similar distance  $d_2$  on  $\mathcal{T}_2(\mathcal{F}_n^m)$ :

$$d_2(\tau_1, \tau_2)$$

$$= \sum_{k=1}^{\infty} 2^{-k} \sup_{(i_1, \dots, i_m) \in \llbracket 1, m \rrbracket^k} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{-1, 1\}^k} \sup_{(j_1, \dots, j_k) \in \{2, \dots, n+1\}^k} |(\tau_1 - \tau_2) \left( (u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_k}^{i_k})^{\epsilon_k} \right)|$$

$$+ \sup_{l=1, \dots, m} \sup_{j=1, \dots, n} \left| (\tau_1 - \tau_2) \left( \left( \frac{u_j^l + 1}{u_j^l - 1} \right)^* \left( \frac{u_j^l + 1}{u_j^l - 1} \right) \right) \right|$$

$$+ \sum_{k=1}^{\infty} 2^{-k} \sup_{(l, i_1, \dots, i_m) \in \llbracket 1, m \rrbracket^{k+1}} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{-1, 1\}^k} \sup_{(j_1, \dots, j_k, j) \in \{2, \dots, n+1\}^{k+1}} |(\tau_1 - \tau_2) \left( \left( \frac{u_j^l + 1}{u_j^l - 1} \right)^* (u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_k}^{i_k})^{\epsilon_k} \right)|.$$

Finally we have two variant spaces  $\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^m)$ , (resp.  $\mathcal{T}_{2,0}^c(\mathcal{F}_n^m * \mathcal{F}_\mu^\nu)$ ) where only the variables in  $\mathcal{F}_{[0,1]}^m$  are associated with  $X_t^i$  with finite second moment in the GNS representation (resp. only variables in the first copy  $\mathcal{F}_n^m$ ). We put e.g. on the first space the distance:

$$d_{2,0}(\tau_1, \tau_2) = d_{0,0}(\tau_1, \tau_2) + \sup_{l=1, \dots, m} \sup_{t \in [0,1]} \left| (\tau_1 - \tau_2) \left( \left( \frac{u_t^l + 1}{u_t^l - 1} \right)^* \left( \frac{u_t^l + 1}{u_t^l - 1} \right) \right) \right|$$

$$+ \sum_{k=1}^{\infty} 2^{-k} \sup_{t \in [0,1]} \sup_{(l, i_1, \dots, i_m) \in \llbracket 1, m \rrbracket^{k+1}} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{-1, 1\}^k} \sup_{(j_1, \dots, j_k) \in (\{2, \dots, \mu+1\} \cup [0, 1])^k} |(\tau_1 - \tau_2) \left( \left( \frac{u_t^l + 1}{u_t^l - 1} \right)^* (u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_k}^{i_k})^{\epsilon_k} \right)|.$$

where

$$d_{0,0}(\tau_1, \tau_2) = \sum_{k=1}^{\infty} 2^{-k} \sup_{(i_1, \dots, i_m) \in \llbracket 1, m \rrbracket^k} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{-1, 1\}^k} \sup_{(j_1, \dots, j_k) \in (\{2, \dots, \mu+1\} \cup [0, 1])^k} |(\tau_1 - \tau_2) \left( (u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_k}^{i_k})^{\epsilon_k} \right)|.$$

The extra unitary variables will be used to generate the relative algebra  $B$  appearing in Theorem A and will be used as variables approximated by non-random matrices. We will only consider any of our considerations concerning classes of convex functions in the next subsection to be uniform over these extra variables. For  $X$  a process in  $L^2(M, \tau)^m$  continuous as before and  $u \in \mathcal{U}(M)^{\mu\nu}$  (we call  $\mathcal{U}(M)$  the set of unitaries), we write  $\tau_{X,u} \in \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu)$  for the joint law and similar notations for laws on  $\mathcal{T}_{2,0}(\mathcal{F}_n^m * \mathcal{F}_\mu^\nu)$ .

**2.2. Some non-commutative continuous convex functions.** We fix  $\mu, \nu \in \mathbb{N}, m \geq 1$ . We will consider two kinds of convex functions on  $\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu)$  bounded from below and with subquadratic growth in the continuous time variables uniformly over the supplementary unitary variables. Of course, this includes the case  $\mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m)$  if  $\nu = \mu = 0$ . To deal with the growth condition, we will need to consider functions depending only on finitely many times. We thus define a subspace of the space of continuous functions  $C^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  defined as :

$$C_{(k)}^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) = \{f \in C^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) : \exists t_1, \dots, t_k \in ]0, 1[$$

$$\exists g \in C^0(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0}), f(\tau) = g(\tau \circ (I_{t_1, \dots, t_k} * Id))\}.$$

For  $f \in C^0_{(k)}(\mathcal{T}_{2,0}^c(\mathcal{F}_\mu^m * \mathcal{F}_\mu^\nu), d_{2,0})$ . we then write  $\mathbf{t}(f)$ ,  $k(f)$  the minimal choice of time sets and index  $k$ .

It is convenient to define for  $\mathbf{t} = (0 < t_1 < t_2 < \dots < t_k)$  the function

$$g_{2,\mathbf{t}}(\tau) = \frac{1}{2} \sum_{l=1}^m \left( \frac{16}{t_1} \tau \left( \left( \frac{u_1^l + 1}{u_1^l - 1} \right)^* \left( \frac{u_1^l + 1}{u_1^l - 1} \right) \right) \right. \\ \left. + \sum_{L=2}^k \frac{16}{t_L - t_{L-1}} \tau \left( \left( \frac{u_L^l + 1}{u_L^l - 1} - \frac{u_{L-1}^l + 1}{u_{L-1}^l - 1} \right)^* \left( \frac{u_L^l + 1}{u_L^l - 1} - \frac{u_{L-1}^l + 1}{u_{L-1}^l - 1} \right) \right) \right).$$

We have  $g_{2,\mathbf{t}} \in C^0(\mathcal{T}_2(\mathcal{F}_k^m), d_2)$ . It corresponds to the potential giving the density of the finite dimensional distribution of hermitian brownian motion.

**Definition 2.1.** A (real valued) continuous function  $g \in C^0(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  is said to be **universally convex** if for any tracial  $W^*$  probability space  $(M, \tau)$ ,  $X \mapsto f(\tau_{X,u})$  is convex on  $(L^2(M, \tau)^{mk})_{sa}$  for any unitaries  $u \in \mathcal{U}(M)^{\mu\nu}$ . It is said **matricially convex** if this holds only for  $M \subset R^\omega$ , some (countable) ultrapower of the hyperfinite  $II_1$  factor  $R$ . A function  $f \in C^0_{(k)}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  is said to be matricially or universally convex if  $f(\tau) = g(\tau \circ (I_{t_1, \dots, t_k} * Id))$  and if so is  $g$ .

For a convex function  $g \in C^0(\mathcal{T}_2(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_2)$ , we may sometimes call, for  $u \in \mathcal{U}(M)^{\mu\nu}$ ,  $g_\tau(u) : X \mapsto f(\tau_{X,u})$  the convex function on  $(L^2(M, \tau)^{mk})_{sa}$ . We can consider its subdifferential, a multivalued map:

$$\partial(g_\tau(u)) : (L^2(M, \tau)^{mk})_{sa} \rightarrow P((L^2(M, \tau)^{mk})_{sa})$$

(see e.g [ET] or [B, ex 2.1.4]) which is defined by:

$$\partial(g_\tau(u))(X) := \{G \in (L^2(M, \tau)^{mk})_{sa} : \forall Y \in L^2(M, \tau)^{mk}, \\ \tau \left( \sum_{i=1}^n \sum_{j=1}^k (G_j^i)^* (Y_j^i - X_j^i) \right) \leq g(\tau_{Y,u}) - g(\tau_{X,u})\}.$$

Note that  $g_\tau(u)$  is continuous on  $L^2(M, \tau)^{mk}$ . Indeed, for any sequence  $X_n \rightarrow X$ , it is easy to see that  $d_2(\tau_{X_n}, \tau_X) \rightarrow 0$ . Thus, from [B, ex 2.3.4],  $\partial(g_\tau(u))$  is a maximal monotone operator so that we will have available the theory of [B].

**Definition 2.2.** We call  $\mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_2)$  the set of universally convex functions  $g \in C^0((\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_2)$ , bounded bellow and **subquadratic** in the sense that there is a  $C > 0$  such that for any  $\tau \in \mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu)$ :

$$(2.1) \quad g(\tau) \leq C \left( 1 + \sum_{l=1}^m \sum_{j=1}^k \tau \left( \left( 4i \frac{u_j^l + 1}{u_j^l - 1} \right)^* \left( 4i \frac{u_j^l + 1}{u_j^l - 1} \right) \right) \right).$$

We call  $\mathcal{E}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  the subset of functions  $g$  of  $\mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  such that there is a constant  $C$  such that for all  $(M, \tau)$  and  $X, Y \in L^2(M, \tau)^{mk}$ ,  $u, v \in \mathcal{U}(M)^{\mu\nu}$ :

$$(2.2) \quad |g(\tau_{X,u}) - g(\tau_{Y,u})| \leq C \left( \sum_{l=1}^m \sum_{j=1}^k \tau((X_j^l - Y_j^l)^*(X_j^l - Y_j^l)) \right)^{1/2} \left( 1 + \sum_{l=1}^m \sum_{j=1}^k \tau((X_j^l)^* X_j^l + (Y_j^l)^* Y_j^l) \right)^{1/2}.$$

$$(2.3) \quad |g(\tau_{X,u}) - g(\tau_{X,v})| \leq C \left( \sum_{l=1}^\mu \sum_{j=1}^\nu \tau((u_j^l - v_j^l)^*(u_j^l - v_j^l)) \right)^{1/2}.$$



and

$$(2.4) \quad |g(\tau_{X+Y,u}) + g(\tau_{X-Y,u}) - 2g(\tau_{X,u})| \leq C \left( \sum_{l=1}^m \sum_{j=1}^k \tau((Y_j^l)^* Y_j^l) \right).$$

and also for some constant  $C(g) > 0$ :

$$(2.5) \quad |g(\tau_1) - g(\tau_2)| \leq C(g) d_{2,0}(\tau_1, \tau_2).$$

We call  $\mathcal{E}_{(k)}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  the set of universally convex functions  $f \in C^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  with  $f(\tau) = g(\tau \circ (I_{t_1, \dots, t_k} * Id))$  for  $g \in \mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  and some  $0 < t_1 < \dots < t_k \leq 1$ .

Finally, we write

$$\mathcal{E}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) = \cup_{k=1}^\infty \mathcal{E}_{(k)}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}),$$

and similarly  $\mathcal{E}_{(k)}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ ,  $\mathcal{E}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ .

In order to prove a Laplace principle [DE] (or a Large deviation principle modulo Bryc's theorem [DZ, Th 4.4.2]), we will need the following lemma to use the variant [DZ, Th 4.4.10]. Recall that a class  $G$  of continuous real valued functions on a topological space  $X$  is said to be *well-separating* if it contains constant functions, is closed by finite pointwise maxima and separates points of  $X$  in the sense that for  $x \neq y \in X$  and  $a, b \in \mathbb{R}$  there exists  $g \in G$  with  $g(x) = a, g(y) = b$ . (note this would say  $-G$  well-separating with the definition of [DZ, Def 4.4.7] but we prefer to discuss convex instead of concave functions.)

For our purpose, we also introduce several classes of explicit functions:

$$\begin{aligned} \mathcal{E}_{reg,\infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \\ = \{f \in \mathcal{E}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) : \exists l \in \mathbb{N}, \exists g, g_1, \dots, g_l \in \mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0}), \\ \exists (D, D_1, \dots, D_l, \lambda_1, \dots, \lambda_l, C_1, \dots, C_l) \in \mathbb{R}^{l+1} \times \mathbb{C}^l \times ]0, \infty[^l, \exists (\epsilon_1^i, \dots, \epsilon_{m_i}^i) \in \{-1, 1\}^{m_i}, \\ \exists ((j_1^i, l_1^i), \dots, (j_{m_i}^i, l_{m_i}^i)) \in (\{1, \dots, k+1\} \times \{1, \dots, m\} \cup \{k+2, \dots, k+\mu+1\} \times \{m+1, \dots, m+\nu\})^{m_i} \\ f(\tau) = g(\tau \circ (I_{t_1, \dots, t_k} * Id)) \text{ and } g(\tau) = D + \left( \max_{i=1, \dots, l} g_i(\tau) \right) \\ \text{and } g_i(\tau) = D_i + C_i \sum_{j=1}^k \sum_{l=1}^m \tau((4i \frac{u_j^l + 1}{u_j^l - 1})^* (4i \frac{u_j^l + 1}{u_j^l - 1})) + \Re(\lambda_i \tau((u_{j_1^i}^{l_1^i})^{\epsilon_1^i} \dots (u_{j_{m_i}^i}^{l_{m_i}^i})^{\epsilon_{m_i}^i})) \geq 1\} \end{aligned}$$

with  $\Re$  the real part to stay within real potentials and another  $p$ -norm variant for  $p \in [1, \infty[$ :

$$\begin{aligned} \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \\ = \{f \in \mathcal{E}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) : \exists l \in \mathbb{N}, \exists g, g_1, \dots, g_l \in \mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0}), \\ \exists (D, D_1, \dots, D_l, \lambda_1, \dots, \lambda_l, C_1, \dots, C_l) \in \mathbb{R}^{l+1} \times \mathbb{C}^l \times ]0, \infty[^l, \exists (\epsilon_1^i, \dots, \epsilon_{m_i}^i) \in \{-1, 1\}^{m_i}, \\ \exists ((j_1^i, l_1^i), \dots, (j_{m_i}^i, l_{m_i}^i)) \in (\{2, \dots, k+1\} \times \{1, \dots, m\} \cup \{k+2, \dots, k+\mu+1\} \times \{m+1, \dots, m+\nu\})^{m_i} \\ \forall \bar{\tau} \in \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), \forall \tau \in \mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu) : \end{aligned}$$

$$\begin{aligned} f(\bar{\tau}) = g(\bar{\tau} \circ (I_{t_1, \dots, t_k} * Id)) \text{ and } g(\tau) = D + \left( \sum_{i=1, \dots, l} (g_i(\tau))^p \right)^{1/p} \\ \text{and } g_i(\tau) = D_i + C_i \sum_{j=1}^k \sum_{l=1}^m \tau((4i \frac{u_j^l + 1}{u_j^l - 1})^* (4i \frac{u_j^l + 1}{u_j^l - 1})) + \Re(\lambda_i \tau((u_{j_1^i}^{l_1^i})^{\epsilon_1^i} \dots (u_{j_{m_i}^i}^{l_{m_i}^i})^{\epsilon_{m_i}^i})) \geq 1\} \end{aligned}$$

We also write  $\mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  the corresponding spaces of functions of the type of  $g$  in the definition above. Note immediately that for any  $G \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ ,  $p \in [1, \infty]$ , there is a constant  $C(G)$  such that (2.5) holds.

**Lemma 2.3.** *The classes  $\mathcal{E}_{reg,\infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , and  $\mathcal{E}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  are well-separating on  $(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ .*

*Proof.* It is obvious that all the spaces contain constant functions and are stable by finite maxima, since convexity is stable by supremum and so is the subquadratic behaviour.

Since  $\mathcal{E}_{reg,\infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \subset \mathcal{E}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , and the first space is also by definition stable by maximum, it only remains to check the separation property on the first space. But since this space is stable by translation (addition of constants) and multiplication by positive numbers, it suffices to find two functions with different values on  $\tau_1 \neq \tau_2$ . But if all the values were the same, we would have (by the cases  $l = 1$   $(C_1, \lambda_1) = (C, \lambda)$  or  $(C, 0)$ ),  $\tau_1((u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_{m_i}}^{i_{m_i}})^{\epsilon_{m_i}}) = \tau_2((u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_{m_i}}^{i_{m_i}})^{\epsilon_{m_i}})$  for all possible choices and thus  $\tau_1 = \tau_2$  since indeed for  $C$  fixed, there is always a  $\lambda$  small enough (real or imaginary) such that  $g(\tau) = C \sum_{j=1}^k \sum_{l=1}^m \tau((4i \frac{u_j^l + 1}{u_j^l - 1}) * (4i \frac{u_j^l + 1}{u_j^l - 1})) + \Re(\lambda \tau((u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_{m_i}}^{i_{m_i}})^{\epsilon_{m_i}}))$  gives  $g \in \mathcal{E}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ .  $\square$

**Definition 2.4.** We call  $\mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0}) \subset \mathcal{E}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  the subset of functions  $g$  such that for all  $I = 1, \dots, m, l = 1, \dots, k$  all  $\epsilon > 0$ , there is  $P_1, \dots, P_L \in \mathcal{F}_k^m * \mathcal{F}_\mu^\nu$ ,  $f_1, \dots, f_L, g_{1,1}, \dots, g_{k,m} \in C^0(\mathcal{T}_{2,0}^c(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  such that for all  $(M, \tau)$  and  $X \in L^2(M, \tau)^{mk}$ ,  $u \in \mathcal{U}(M)^{\mu\nu}$  we have the approximation:

$$(2.6) \quad \|\nabla_{X_j^{(I)}} g_\tau(u)(X) - \sum_{i=1}^L P_i(u(X), u) f_i(\tau_{X,u}) + \sum_{i=1}^m \sum_{j=1}^k X_j^{(i)} g_{j,i}(\tau_{X,u})\|_2 \leq C\epsilon.$$

We then define  $\mathcal{E}_{app(k)}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  and  $\mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  as before.

Note that we will see later (proposition 2.11) that the assumption implies  $g_\tau$  differentiable so that the subddifferential  $\partial g_\tau(X) = \{(\nabla_{X_j^{(I)}} g_\tau(X))_{j,I}\}$  is singleton valued.

Clearly  $\mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \subset \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  for  $p \in [2, \infty[$  (cf the proof of lemma 5.3 below for an explicit computation of derivatives making that point more explicit).

**2.3. Malliavin calculus and distributional Clark-Ocone formula.** We give in this subsection preliminaries on Malliavin calculus as background to understand the statements from [Us14] that we will use.

Let  $\mathbb{W} \subset C^0([0, 1], \mathbb{R}^d)$  the set of continuous paths starting at 0 and  $\gamma$  Wiener measure on it.  $B$  will be the canonical coordinate process. As usual  $\mathbb{H} \subset \mathbb{W}$  is the Calderon-Martin space of functions  $U_t = \int_0^t u_s ds$  with  $\|U\|_{\mathbb{H}} = \int_0^1 |u_s|^2 ds < \infty$ . We then write  $u_s = \dot{U}_s$ .

Since the translations of  $\gamma$  with the elements of  $\mathbb{H}$  induce measures equivalent to  $\gamma$ , the Gâteaux derivative in  $\mathbb{H}$  direction of the random variables is a closable operator on  $L^p(\gamma)$ -spaces and this closure will be denoted by  $\nabla : L^p(\mathbb{W}, \gamma) \rightarrow L^p((\mathbb{W}, \gamma : \mathbb{H}))$  (cf. e.g. [Us95, Us10, N06]). It is also useful to point out the explicit formula for  $f \in C^1(\mathbb{R}^{kd})$  a  $C^1$  function,  $t_1 < t_2 < \dots < t_k \in [0, 1]$  and  $F = f(B_{t_1}, \dots, B_{t_k})$ , we have (if  $h \in \mathbb{H}$  is the coordinate and  $d_i f$  is the partial differential in the directions of the  $i$ -th  $d$ -uple of variables,  $\nabla_i f$  the corresponding gradient with  $d_i f(x).h = \langle \nabla_i f(x), h \rangle$ ):

$$\nabla F = \sum_{i=1}^n (d_i f)(B_{t_1}, \dots, B_{t_k}).h_{t_i}$$

It is then common to write for  $h \in \mathbb{H}$ :

$$\langle \nabla F, h \rangle = \nabla_h F = \int_0^1 \dot{h}_t D_t F dt$$

where  $D_t F$  is the Lebesgue density of the process  $\nabla F$  seen in  $L^p(\Omega, \gamma : \mathbb{H})$  defined  $d\gamma \otimes dt$  almost surely. Note also that  $\nabla_h : L^p(\mathbb{W}, \gamma) \rightarrow L^2(\mathbb{W}, \gamma)$  is also a closable operator for  $h \in \mathbb{H}$ .

Especially for  $F = f(B_{t_1}, \dots, B_{t_k})$  as above, it is easy to see that :

$$(2.7) \quad D_t F = \sum_{i=1}^n 1_{[0, t_i]}(t) (\nabla_i f)(B_{t_1}, \dots, B_{t_k}).$$

The corresponding Sobolev spaces of (the equivalence classes of) real random variables will be denoted as  $\mathbb{D}_{p,k}$ , where  $k \in \mathbb{N}$  is the order of differentiability and  $p > 1$  is the order of integrability. If the random variables are with values in some separable Hilbert space, say  $\Phi$ , then we can define similarly the corresponding Sobolev spaces and they are denoted as  $\mathbb{D}_{p,k}(\Phi)$ ,  $p > 1$ ,  $k \in \mathbb{N}$ . Since  $\nabla : \mathbb{D}_{p,k} \rightarrow \mathbb{D}_{p,k-1}(H)$  is a continuous and linear operator its adjoint is a well-defined operator which we represent by  $\delta$ .  $\delta$  coincides with the Itô integral of the Lebesgue density of the adapted elements of  $\mathbb{D}_{p,k}(H)$  (cf. [Us95, Us10, N06]).

We denote by  $\mathbb{D}_{p,k}^a(\mathbb{H})$  the subspace defined by

$$\mathbb{D}_{p,k}^a(\mathbb{H}) = \{\xi \in \mathbb{D}_{p,k}(\mathbb{H}) : \dot{\xi} \text{ is adapted}\}$$

for  $p > 1$ ,  $k \in \mathbb{N}$ , for  $p = 2$ ,  $k = 0$ , we shall write  $L_a^2(\mu, \mathbb{H})$ .

To use the results of [Us14], we will need supplementary background from [Us87]. We consider  $\mathbb{D}(\Phi) = \bigcap_{p,k} \mathbb{D}_{p,k}(\Phi)$ ,  $\mathbb{D}^a(\mathbb{H}) = \bigcap_{p,k} \mathbb{D}_{p,k}^a(\mathbb{H})$  with projective limit topology and the continuous duals  $\mathbb{D}'(\Phi)$ ,  $(\mathbb{D}^a(\mathbb{H}))'$  (if  $\Phi = \mathbb{C}$  we write only  $\mathbb{D}$ ,  $\mathbb{D}'$ .) We also let  $\mathbb{D}_0 = \overline{\{\psi - \langle \psi, 1 \rangle, \psi \in \mathbb{D}\}}$  In [Us87, Corol II.1], it is shown that  $J : \mathbb{D}^a(\mathbb{H}) \rightarrow \mathbb{D}_0$  defined by stochastic integration

$$J(\xi) = \int_0^1 \dot{\xi}_s dB_s$$

is continuous and has a continuous inverse  $\partial_B : \mathbb{D}_0 \rightarrow \mathbb{D}^a(\mathbb{H})$  which enables to express Clark-Ocone's formula as

$$\psi = \langle \psi, 1 \rangle + J(\partial_B(\psi - \langle \psi, 1 \rangle)).$$

Then, the adjoint of  $J$  extends  $\partial_B$ , and the adjoint of  $\partial_B$  extends  $J$  so that if we still call in the same way the extensions, the previous formula holds for  $\psi \in \mathbb{D}'$  [Us87, Prop II.2]. For  $\xi \in \mathbb{D}(\mathbb{H})$ , if  $\pi\xi$  is an  $\mathbb{H}$ -valued process with Lebesgue derivative  $E(\dot{\xi}_s | \mathcal{F}_s)$ . Then [Us87, Prop III.1],  $\pi : \mathbb{D}(\mathbb{H}) \rightarrow \mathbb{D}^a(\mathbb{H})$  is continuous and has a unique continuous extension  $\hat{\pi} : \mathbb{D}'(\mathbb{H}) \rightarrow (\mathbb{D}^a(\mathbb{H}))'$  [Us87, Prop III.1] which coincides with the restriction map on  $\mathbb{D}^a(\mathbb{H}) \subset \mathbb{D}(\mathbb{H})$ . Then on  $\mathbb{D}'_0 = \{\psi - \langle \psi, 1 \rangle, \psi \in \mathbb{D}'\}$ , [Us87, Th IV.1] gives the representation

$$\partial_B = \hat{\pi} \nabla.$$

Especially, if  $F \in L^{1+\epsilon}(\gamma)$ ,  $\epsilon > 0$ , we can consider  $\nabla F \in \mathbb{D}_{1+\epsilon,-1}(\mathbb{H})$  and  $\partial_B(F) = \hat{\pi} \nabla(F) \in L^{1+\epsilon}(\gamma; \mathbb{H})$  using Buchholder-Davis-Gundy inequality giving the equivalence of norms induced by  $J : L_{ad}^{1+\epsilon}(\gamma; \mathbb{H}) \rightarrow L^{1+\epsilon}(\gamma)$ . In that case, we will write

$$[\hat{\pi} \nabla(F)]_s = E(D_s F | \mathcal{F}_s).$$

By the extension properties mentioned before, this expression coincides when  $F = f(B_{t_1}, \dots, B_{t_k})$  with the previous expression if  $f$  is  $C^1$ . We will thus be able to use results in [Us14] expressed in terms of  $[\hat{\pi} \nabla(F)]_s$  in using classical Malliavin calculus formulas they extend.

Finally, a measurable function  $f : \mathbb{W} \rightarrow \mathbb{R} \cup \{+\infty\}$  is called  $\alpha$ -convex,  $\alpha \in \mathbb{R}$ , if the map

$$h \rightarrow f(x+h) + \frac{\alpha}{2} |h|_H^2 = F(x, h)$$

is convex on the Cameron-Martin space  $\mathbb{H}$  with values in  $L^0(\mu)$ . Note that this notion is compatible with the  $\mu$ -equivalence classes of random variables thanks to the Cameron-Martin theorem (cf. [FU]).

**2.4. Hermitian brownian motion and its exponential tightness.** For  $(N, m, \mu, \nu) \in (\mathbb{N}^*)^4$ , and  $d = N^2 m$ , we write  $\mathbb{W}_{sa, N} \subset C^0([0, 1], \mathbb{R}^d) = C^0([0, 1], (M_N(\mathbb{C})_{sa})^m)$  the Wiener space for paths on  $m$ -tuples of hermitian matrices in  $M_N(\mathbb{C})_{sa} \simeq \mathbb{R}^{N^2}$  ( $m$  is fixed throughout the paper and does not appear in the notation). If  $\gamma$  is the law of the standard brownian measure making  $B_t$  into an hermitian brownian motion, we write  $\gamma_{sa, N, m} = \gamma_N$  the law of  $\frac{B_t}{\sqrt{N}}$ . We define a random state  $\hat{\sigma}^N \in \mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m)$  by  $\hat{\sigma}^N = \tau_{H^N}$  with  $H^N = (\frac{B_t}{\sqrt{N}})_{t \in [0,1]}$ . Note that the normalisation is as usual made to insure :

$$E(\hat{\sigma}_N(4i \frac{u_t^l + 1}{u_t^l - 1})^* (4i \frac{u_t^l + 1}{u_t^l - 1})) = E(\frac{1}{N} \text{Tr}((\frac{B_t^l}{\sqrt{N}})^2)) = \frac{tN^2}{N^2} = t.$$

If  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$  is a bunch of (deterministic) unitary matrices, we call

$$\hat{\sigma}_{\Upsilon}^N = \tau_{H^N, \Upsilon} \in \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_{\mu}^{\nu}).$$

We will be interested in large deviation results for this  $\hat{\sigma}_{\Upsilon}^N$  (this includes if  $\mu = \nu = 0$  the case  $\hat{\sigma}^N$  considered in [BCG]).

We can also define the image of the Gaussian Unitary Ensemble in Unitary variables as above  $\mathfrak{G}_N = \tau_{G^N} \in \mathcal{T}_2(\mathcal{F}_1^m)$ , with  $G^N = (\frac{B_1}{\sqrt{N}})$ .

The only statement we will use from [BCG] to prove our large deviation principle is their lemma 5.4 (or rather actually a slight improvement, consequence of their proof) giving exponential tightness of  $\hat{\sigma}_{\Upsilon}^N$  in  $\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_{\mu}^{\nu})$ . Let us recall the appropriate notation. For any  $g : [0, 1] \rightarrow \mathbb{R}^+$  with  $\lim_{x \rightarrow 0} g(x) = 0$ . We let :

$$\begin{aligned} K_g &= \left\{ \tau \in \mathcal{T}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_{\mu}^{\nu}) : \forall s \leq t \in [0, 1] \max_{i=1 \dots m} \tau(|u_t^i - u_s^i|^2) \leq g(t-s) \right\}, \\ K_{g,2} &= \left\{ \tau \in \mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_{\mu}^{\nu}) : \forall s \leq t \in [0, 1] \max_{i=1 \dots m} \tau(|4i \frac{u_t^l + 1}{u_t^l - 1} - 4i \frac{u_s^l + 1}{u_s^l - 1}|^2) \leq g(t-s) \right\} \subset K_g, \\ \Gamma_L &= \left\{ \tau \in \mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_{\mu}^{\nu}) : \forall t \in [0, 1] \max_{i=1 \dots m} \tau(|4i \frac{u_t^l + 1}{u_t^l - 1}|^2) \leq L \right\}. \end{aligned}$$

**Lemma 2.5.** *For any  $L > 0, g$  as above,  $K_{g,2} \cap \Gamma_L$  is a compact set of  $\mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_{\mu}^{\nu})$ .*

*Proof.* First, the set is obviously closed. Since  $K_{g,2} \subset K_g$ , one can argue as in [BCG, lemma 2.1] to check the set is precompact for  $d$ . Similarly as in their proof, the family of maps  $t \mapsto \tau((4i \frac{u_t^l + 1}{u_t^l - 1})^* (4i \frac{u_t^l + 1}{u_t^l - 1}))$ ,  $\tau \in K_{g,2} \cap \Gamma_L$  is equicontinuous and pointwise bounded, thus the result follows from Arzela-Ascoli theorem. This gives precompactness for  $d_{2,0}$ .  $\square$

From the proof of their result, one readily deduces ( since  $\Upsilon_N$  does not appear in the sets above):

**Lemma 2.6** (Lemma 5.4 in [BCG]). *For any sequence  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$ ,  $\hat{\sigma}_{\Upsilon_N}^N$  is exponentially tight since :*

$$\limsup_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N^2} \log P \left( \hat{\sigma}_{\Upsilon_N}^N \in (K_{L\sqrt{\cdot}, 2} \cap \Gamma_L)^c \right) = -\infty.$$

**2.5. Concentration of measure for Random matrices.** We will need two kinds of concentration of measure results, one for the lower bound (usual LDP upper bound) and one for the upper bound, the first one being mostly used to obtain an appropriate set up to be able to apply the second one.

The first result uses convexity of a potential and Brascamp-Lieb inequality. It was first used in the proof of [GM, Theorem 3.4]. We follow their method and only give the proof for the reader's convenience. Following the probabilistic tradition, we state the almost sure result, but the interesting for us with our use of ultraproducts techniques is the uniform integrability like bound (2.8). In a second part, we also include a concentration property coming from logarithmic Sobolev inequality (see e.g. [AGZ, Th 4.4.17]).

For brevity, we write  $\mathcal{C}_k^m = \mathbb{C}\langle X_1^1, \dots, X_1^m, X_2^1, \dots, X_k^m \rangle$  the algebra of non-commutative polynomials in selfadjoint variables and

$$\mathcal{C}_{k,\mu}^{m,\nu} = \mathbb{C}\langle X_1^1, \dots, X_1^m, X_2^1, \dots, X_k^m, u_1^1, \dots, u_1^\nu, u_2^1, \dots, u_\mu^\nu \rangle,$$

the algebra of non-commutative polynomials in the same selfadjoint variables and supplementary unitary variables.

**Proposition 2.7.** *Let  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$  a sequence of unitary matrices. Let  $g \in \mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  and  $\mathbf{t} = (0 < t_1 < t_2 < \dots < t_k)$  and consider the probability on  $(M_N(\mathbb{C})_{sa})^{km}$  given (for some normalization constant  $Z_{g,\mathbf{t},N}$ ) by :*

$$\mu_{g,\mathbf{t},N}(dx) = \frac{1}{Z_{g,\mathbf{t},N}} e^{-N^2 g(\tau_x, \Upsilon_N) - N^2 g_{2,\mathbf{t}}(\tau_x)} dLeb_{(M_N(\mathbb{C})_{sa})^{km}}(dx)$$

Let  $A_1^N, \dots, A_k^N = (A_{k,1}^N, \dots, A_{k,m}^N)$  of law  $\mu_{g,\mathbf{t},N}$  (on a same probability space), we have a constant  $C > 0$  such that a.s.:

$$\limsup_{N \rightarrow \infty} \max_i \|A_i^N\|_\infty \leq C,$$

and for  $K \in \mathbb{N}^*$

$$(2.8) \quad \limsup_{N \rightarrow \infty} E_{\mu_{g,\mathbf{t},N}}(1_{\{\|A_{i,l}^N\|_\infty \geq C\}} \frac{1}{N} Tr((A_{i,l}^N)^{2K})) = 0.$$

Moreover, for any non-commutative polynomial  $P \in \mathcal{C}_{k,\mu}^{m,\nu} \otimes_{alg} \mathcal{C}_{k,\mu}^{m,\nu}$

$$\lim_{N \rightarrow \infty} \left| E_{\mu_{g,\mathbf{t},N}} \left( \frac{1}{N^2} (Tr \otimes Tr)(P(A_1, \dots, A_k)) - \frac{1}{N^2} [(E_{\mu_{g,\mathbf{t},N}} \circ Tr) \otimes (E_{\mu_{g,\mathbf{t},N}} \circ Tr)](P) \right) \right| = 0.$$

*Proof.* We follow the beginning of the proof of [GM, Theorem 3.4]. By Brascamp-Lieb inequality [Ha, theorem 1.1] (and [S, Th 2 p 709] for the second inequality), we have (if  $E(A_{i,l}^N) = E_{\mu_{g,\mathbf{t},N}}(A_{i,l}^N)$  entrywise)

$$\mu_{g,\mathbf{t},N} \left( \frac{1}{N} Tr((A_{i,l}^N - E(A_{i,l}^N))^{2k}) \right) \leq \mu_{0,\mathbf{t},N} \left( \frac{1}{N} Tr((A_{i,l}^N)^{2k}) \right) \leq C 4^k t_i^{2k}, \quad k = \sqrt{N},$$

and for  $k \geq K$  (using also in the second line also [MN, Th 7.5] and Jensen's inequality):

$$(2.9) \quad \begin{aligned} \mu_{g,\mathbf{t},N} \left( \left( \frac{1}{N} Tr((A_{i,l}^N - E(A_{i,l}^N))^{2k}) \right)^{1+K/k} \right) &\leq \mu_{0,\mathbf{t},N} \left( \left( \frac{1}{N} Tr((A_{i,l}^N)^{2k}) \right)^{1+K/k} \right) \\ &\leq \left[ \mu_{0,\mathbf{t},N} \left( \left( \frac{1}{N} Tr((A_{i,l}^N)^{2k}) \right)^2 \right) \right]^{1/2+K/2k} \leq C_K 4^{k+K} t_i^{2k+2K}. \end{aligned}$$

And thus by Markov inequality, one gets:

$$\begin{aligned} \mu_{g,\mathbf{t},N}(\|A_{i,l}^N - E(A_{i,l}^N)\|_\infty \geq 3t_i) &\leq \mu_{g,\mathbf{t},N} \left( \frac{1}{N} Tr((A_{i,l}^N - E(A_{i,l}^N))^{2\sqrt{N}}) \geq \frac{1}{N} (3t_i)^{2\sqrt{N}} \right) \\ &\leq C N \left( \frac{2}{3} \right)^{2\sqrt{N}}. \end{aligned}$$



Moreover, we also have the uniform integrability type bound in using the same Markov inequality type argument for  $N$  large enough:

$$\begin{aligned}
& E_{\mu_{g,t,N}}(1_{\|A_{i,l}^N - E(A_{i,l}^N)\|_\infty \geq 3t_i} \frac{1}{N} \text{Tr}((A_{i,l}^N)^{2K})) \\
& \leq E_{\mu_{g,t,N}}(1_{\frac{1}{N} \text{Tr}((A_{i,l}^N - E(A_{i,l}^N))^{2\sqrt{N}})} \geq \frac{1}{N} (3t_i)^{2\sqrt{N}}} \left( \frac{2^K}{N} \text{Tr}(((A_{i,l}^N - E(A_{i,l}^N))^{2K})) + \frac{2^K}{N} \text{Tr}((E(A_{i,l}^N))^{2K}) \right) \\
& \leq \frac{2^K}{N} \text{Tr}((E(A_{i,l}^N))^{2K}) C N \left( \frac{2}{3} \right)^{2\sqrt{N}} \\
& + \frac{2^K N}{(3t_i)^{2\sqrt{N}}} E_{\mu_{g,t,N}} \left( \frac{1}{N} \text{Tr}((A_{i,l}^N - E(A_{i,l}^N))^{2\sqrt{N}}) \frac{1}{N} \text{Tr}((A_{i,l}^N - E(A_{i,l}^N))^{2K}) \right) \\
& \leq \frac{2^K}{N} \text{Tr}((E(A_{i,l}^N))^{2K}) C N \left( \frac{2}{3} \right)^{2\sqrt{N}} \\
& + \frac{2^K N}{(3t_i)^{2\sqrt{N}}} E_{\mu_{g,t,N}} \left( \left( \frac{1}{N} \text{Tr}((A_{i,l}^N - E(A_{i,l}^N))^{2\sqrt{N}}) \right)^{1+K/\sqrt{N}} \right) \\
& \leq \frac{2^K}{N} \text{Tr}((E(A_{i,l}^N))^{2K}) C N \left( \frac{2}{3} \right)^{2\sqrt{N}} + \frac{2^K N}{(3t_i)^{2\sqrt{N}}} C K 4^{\sqrt{N}+K} t_i^{2\sqrt{N}+2K}.
\end{aligned}$$

where the next-to-last inequality comes from Hölder inequality for the normalized trace and the last inequality comes from (2.9). But by unitary invariance  $E(A_{i,l}^N) = E(\frac{1}{N} \text{Tr}(A_{i,l}^N)) I d_N$  so that  $\|E(A_{i,l}^N)\|_\infty = |E(\frac{1}{N} \text{Tr}(A_{i,l}^N))|$ .

First note that from the subquadratic growth condition (2.1), applied to  $g$ , one deduces

$$\begin{aligned}
\frac{Z_{g,t}}{Z_{0,t}} & \geq \frac{1}{Z_{0,t}} \int e^{-N^2(C(1+\sum_{l=1}^m \sum_{j=1}^k \tau_x((x_j^l)^* x_j^l) - N^2 g_{2,t}(\tau_x)) d\text{Leb}_{(M_N(\mathbb{C})_{sa})^{km}}(dx)} \\
& \geq \exp \left( - \int N^2(C(1 + \sum_{l=1}^m \sum_{j=1}^k \tau_x((x_j^l)^* x_j^l)) \frac{1}{Z_{0,t}} e^{-N^2 g_{2,t}(\tau_x)} d\text{Leb}_{(M_N(\mathbb{C})_{sa})^{km}}(dx) \right) \\
& = e^{-N^2 C(1+m \sum_{j=1}^k t_j)} =: e^{-N^2 D}
\end{aligned}$$

where the second inequality comes from Jensen's inequality. Now using that  $g$  is bounded below by  $\sup(-g)$ , one deduces from Markov's inequality for any  $y > 0, \lambda > 0$ :

$$\begin{aligned}
\mu_{g,t,N}(|\frac{1}{N} \text{Tr}(A_{i,l}^N)| \geq y) & = \mu_{g,t,N}(e^{\lambda N^2 |\frac{1}{N} \text{Tr}(A_{i,l}^N)|} \geq e^{\lambda N^2 y}) \\
& \leq e^{-\lambda N^2 y + (D - \sup(-g)) N^2} \frac{1}{Z_{0,t}} \int e^{\lambda N^2 |\frac{1}{N} \text{Tr}(A_{i,l}^N)| - N^2 g_{2,t}(\tau_x)} d\text{Leb}_{(M_N(\mathbb{C})_{sa})^{km}}(dx) \\
& \leq e^{-\lambda N^2 y + (D - \sup(-g)) N^2} \frac{1}{Z_{0,t}} \int e^{-\lambda N^2 \frac{1}{N} \text{Tr}(A_{i,l}^N) - N^2 g_{2,t}(\tau_x)} d\text{Leb}_{(M_N(\mathbb{C})_{sa})^{km}}(dx) \\
& + e^{-\lambda N^2 y + (D - \sup(-g)) N^2} \frac{1}{Z_{0,t}} \int e^{\lambda N^2 \frac{1}{N} \text{Tr}(A_{i,l}^N) - N^2 g_{2,t}(\tau_x)} d\text{Leb}_{(M_N(\mathbb{C})_{sa})^{km}}(dx) \\
& \leq 2e^{-\lambda N^2 y + (D - \sup(-g)) N^2 + \frac{N^2}{2} \lambda^2 t_i}
\end{aligned}$$

Thus optimizing in  $\lambda = y/t_i$ , one gets a bound by  $2e^{AN^2 - \frac{y^2 N^2}{2t_i}}$  for  $A = (D - \sup(-g))$ . It is pertinent to cut integrals at  $y = \sqrt{2t_i A}$  in order to get:

$$\mu_{g,t,N}(|\frac{1}{N} \text{Tr}(A_{i,l}^N)|) \leq \sqrt{2t_i A} + 2e^{AN^2} \int_{\sqrt{2t_i A}}^\infty dy e^{-\frac{\sqrt{A} N^2 y}{\sqrt{2t_i}}} \leq \sqrt{2t_i} \left( \sqrt{A} + 2 \frac{1}{\sqrt{A}} \right).$$

From Borel-Cantelli's lemma, this concludes to the almost sure statement with  $C = \sqrt{2t_i} \left( \sqrt{A} + 2 \frac{1}{\sqrt{A}} + 3 \right)$ . The same  $C$  works for the second statement.

For the supplementary statement, it clearly suffices to check for any non-commutative polynomial  $Q \in \mathcal{C}_{k,\mu}^{m,\nu}$  :

$$(2.10) \quad \lim_{N \rightarrow \infty} \frac{1}{N} E_{\mu_{g,t,N}} (|Tr(Q) - (E_{\mu_{g,t,N}} \circ Tr)(Q)|) = 0.$$

If  $x \mapsto g(\tau_{x,v_N})$  were  $C^2$  on any matrix spaces, we could deduce that from [AGZ, Theorem 4.4.17] which is based on Bakry-Emery criterion. Instead, we use [BL, Proposition 3.1] which only uses convexity. Let  $c = \max(t_1, t_2 - t_1, \dots, t_k - t_{k-1}) > 0$ , then  $x \mapsto N^2 g_{2,t}(\tau_x)$  has second derivative bounded below by  $N/c > 0$  and thus satisfy (3.1) in [BL] (with euclidean norm), thus adding a convex potential so does the potential for  $\mu_{g,t,N}$ . Thus, from their proposition 3.1,  $\mu_{g,t,N}$  satisfies logarithmic Sobolev inequality with constant  $c/N$  and thus the Poincaré inequality with constant  $m = N/c$ , (in the sense of [AGZ, Definition 4.4.2], see their [BL, (3.5)]), namely if  $H_{P,N} = \frac{1}{N} Tr(Q)$ :

$$E_{\mu_{g,t,N}} (|H_{P,N} - E_{\mu_{g,t,N}}(H_{P,N})|^2) \leq \frac{c}{N^2} E \left( \sum_{i,l} \frac{1}{N} Tr((\mathcal{D}_i^l P)^* (\mathcal{D}_i^l P)) \right)$$

and since from our previous result the written expectation has bounded limsup, using  $C$  as almost sure bound of our variables, one gets the result.  $\square$

Our second concentration result is a variant adapted to our context of [BD13, lemma 6.1]. Recall that we call  $\mathcal{S}_R^m$  the convex set of tracial states on the universal  $C^*$  algebra free product  $\star_{i=1}^n C^0([-R, R])$ . The key fact at the basis of our concentration result is that  $\mathcal{S}_R^m$  is a Poulsen Simplex in the sense of [LOS]. This property has been proved in [D08, Corollary 5] using free entropy techniques.

We will need a variant for another simplex. We call  $\mathcal{S}_R^m * \mathcal{T}(\mathcal{F}_\mu^\nu)$  the convex set of tracial states on the universal free product  $\star_{i=1}^n C^0([-R, R]) * \mathcal{F}_\mu^\nu$ . Mixing the quoted result with the unitary variant (similar to [DDM, lemma 5.2 and Th 5.3]) one obtains:

**Lemma 2.8.** *If  $m + \mu\nu \geq 2, m \geq 1$ , then  $\mathcal{S}_R^m * \mathcal{T}(\mathcal{F}_\mu^\nu)$  is the Poulsen Simplex.*

*Proof.* The only potentially non-well known statement to check is that the extreme points are dense. If  $\mu\nu = 0$  use [D08, Corollary 5], thus assume  $\mu\nu \geq 1$ . For  $X_1, \dots, X_m, u_1^1, \dots, u_\mu^\nu$  variables in the GNS representation of a state. If  $m \geq 2$  consider  $Y_{i,t} = \frac{R(X_i + tS_i)}{R+2t}$ , with  $S_i$  free semi-circular variables free from  $X_1, \dots, X_m, u_1^1, \dots, u_\mu^\nu$ , as in [D08, Corollary 5], then use [V5] to get  $\Phi^*(Y_{1,t}, \dots, Y_{m,t} : W^*(u_1^1, \dots, u_\mu^\nu)) < \infty$  then since  $Y_{2,t}$  has finite entropy [V5], it is diffuse, and using [D08, Th 4],  $W^*(Y_{1,t}, \dots, Y_{m,t}, u_1^1, \dots, u_\mu^\nu)$  is a factor thus correspond to an extremal state in  $\mathcal{S}_R^m * \mathcal{T}(\mathcal{F}_\mu^\nu)$ . If  $m = 1$ , approximate  $u_1^1$  by a diffuse random variable  $v_t$  and concludes in the same way with  $Y_{1,t}, v_t, u_2^1, \dots, u_\mu^\nu$ .  $\square$

We then quote a variant of [BD13, Corollary 5.4].

The only change is that we consider for  $\tau \in \mathcal{S}_R^m * \mathcal{T}(\mathcal{F}_\mu^\nu)$  another neighbourhood basis of the weak- $*$  topology. We call  $U_{\epsilon,K}(\tau)$  the set of tracial states  $\sigma$  such that for all  $k \leq K$ ,  $((j_1, i_1), \dots, (j_m, i_m)) \in ([1, m] \times \{0\} \cup \{1, \dots, \mu\} \times \{1, \dots, \nu\})^k$   $(\epsilon_1, \dots, \epsilon_m) \in (\{-1, 1\})^k$ , we have

$$|(\sigma - \tau)((u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_k}^{i_k})^{\epsilon_k})| \leq \epsilon,$$

where  $u_j^0 = u(X_j)$  is obtained from the canonical variable  $X_j = X_j(\tau), j = 1, \dots, m$  in the GNS representation of  $L^2(\tau)$  as in subsection 2.1, and  $u(\tau) = (u_j^i)_{(j,i) \in \{2, \dots, \mu+1\} \times \{1, \dots, \nu\}}$  the corresponding unitary variable in the GNS representation. We define  $V_{\epsilon,K}(\tau)$  as in [BD13] in considering instead ordinary monomials in variables  $X_j$  and  $(u_j^i)_{(j,i) \in \{1, \dots, \mu\} \times \{1, \dots, \nu\}}$  of order less than  $K$ .

This defines a map  $X(\tau)$  and recall we also defined  $\tau_X$  as a tracial state on  $\mathcal{F}_1^m * \mathcal{T}(\mathcal{F}_\mu^\nu)$  in subsection 2.1. Note that the map  $\tau \mapsto \tau_{X(\tau), u(\tau)}$  induces a homeomorphism for the weak- $*$  topology to the topology given by  $d_{0,0}$  or  $d_{2,0}$  which is equivalent on the image since  $\frac{1}{u_j - 1} =$

$\frac{1}{(u_j+1)^{-2}}$  has a power expansion since  $\|u_j + 1\| \leq 2\frac{R}{\sqrt{R^2+16}} < 2$ . Note that the homeomorphism statement has a similar proof, the topology of  $d_{2,0}$  is clearly weaker than the image of weak-\* topology since  $u_j \in \star_{i=1}^n C^0([-R, R])$  and stronger in reasoning as above.

We thus deduce from the same proof as [BD13, Corollary 5.4]:

**Lemma 2.9.** *Let  $\tau$  be an extremal state in  $\mathcal{S}_R^n \star \mathcal{T}(\mathcal{F}_\mu^\nu)$  with  $m + \mu\nu \geq 2, m \geq 1$ , and  $\epsilon > 0$ . For any  $\eta > 0$ , there exists a self adjoint polynomial*

$$Q_\eta \in \mathbb{C}\langle u(X_1), u(X_1)^*, \dots, u(X_n), u(X_n)^*, (u_j^i)_{(j,i) \in \{2, \dots, \mu+1\} \times \{1, \dots, \nu\}} \rangle$$

such that for every  $\sigma \in \mathcal{S}_R^n$  one has

$$\tau(Q_\eta) > \sigma(Q_\eta) - \eta$$

and for all  $\sigma \notin V_{\epsilon, K}(\tau)$  (resp.  $\sigma \notin U_{\epsilon, K}(\tau)$ ) one has

$$\sigma(Q_\eta) < \tau(Q_\eta) - 1.$$

We deduce the concentration of measure in the form we need it.

**Proposition 2.10.** *If  $\tau$  is an extremal state in  $\mathcal{S}_R^m \star \mathcal{T}(\mathcal{F}_\mu^\nu)$  with  $m + \mu\nu \geq 2, m \geq 1$ , then for any  $\epsilon > 0$  there exists  $\eta > 0$ , such that for all  $N \in \mathbb{N}^*$ , for any probability measure  $\mu$  law of  $X \in (M_N(\mathbb{C}))_{sa}^n$  supported on the ball of operator norm  $R$  and  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$ , if  $d_{2,0}(E_\mu \circ \tau, \Upsilon_N, \tau_{X(\tau), u(\tau)}) \leq \eta$  then*

$$E_\mu(d_{2,0}(\tau, \Upsilon_N, \tau_{X(\tau), u(\tau)})) \leq \epsilon.$$

*Proof.* Since  $d_{2,0}$  and  $d_{0,0}$  give equivalent topologies on the image of  $\mathcal{S}_R^m \star \mathcal{T}(\mathcal{F}_\mu^\nu)$  by  $\tau_{X(\cdot), u(\cdot)}$  as explained before, it is easy to see that it suffices to prove the corresponding statement for  $d_{0,0}$  instead of  $d_{2,0}$ . Since the distance  $d_{0,0}$  is bounded by 1, it suffices to prove that with probability greater than  $1 - \epsilon/2$  we have  $d(\tau, \Upsilon_N, \tau_{X(\tau), u(\tau)}) \leq \epsilon/2$  and for that, it suffices to bound sufficiently many moments with probability greater than  $1 - \epsilon/2$ . The conclusion follows from our previous lemma as in [BD13, lemma 6.1].  $\square$

**2.6. A lipschitzness criterion for directional derivatives and Lax-Hopf-Yosida semigroup.** It is well-known (see e.g. [FS]) that it is easier to estimate difference quotients of optimal control problems than derivatives. This is why the relation between regularity bounds of difference quotients (or so-called higher order modulus of continuity) and derivatives will be crucial for us. A first idea would be to use a general kind of spaces between Hölder-Zygmund spaces and Besov spaces, the so-called Nikol'skiĭ Spaces (see [KJF, section 8.2] and [N75]), but most of the available results are strongly dimension dependent. Looking for a dimension independent result, we will rather rely on the following well-known result in convex analysis [HP]. In recent terminology, a function both para-convex and para-concave is Gâteaux differentiable with Lipschitz derivative (especially  $f$  is  $C^1$  and thus Fréchet-differentiable). Of course, as usual, for  $f : H \rightarrow \mathbb{R}$ , we look at  $df : H \rightarrow H' \simeq H$  so that we write  $\nabla f(x) \in H$  the vector corresponding to  $df = \langle \nabla f, \cdot \rangle$ .

**Proposition 2.11.** *[2.2.1 in [HP]] Let  $f : H \rightarrow \mathbb{R}$  be a function on a Hilbert space  $H$  with  $\alpha > 0$  such that both  $\frac{\alpha}{2}\|\cdot\|^2 - f$  and  $\frac{\alpha}{2}\|\cdot\|^2 + f$  are convex, then  $f$  is Gâteaux differentiable on  $H$  and we have*

$$(2.11) \quad \|\nabla f(x) - \nabla f(y)\| \leq \alpha\|x - y\|.$$

We will also need a local version for instance proved in finite dimension in [E, lemma 2.1]:

**Proposition 2.12.** *Let  $H = \mathbb{R}^n$  with standard euclidean norm and  $x \in \mathbb{R}^n, \delta > 0, f : B(x, \delta) \rightarrow \mathbb{R}$  be a function defined on the ball centered at  $x$  with  $\alpha > 0$  such that both  $\frac{\alpha}{2}\|\cdot\|^2 - f$  and  $\frac{\alpha}{2}\|\cdot\|^2 + f$  are convex on  $B(x, \delta)$ , then  $f$  is Gâteaux differentiable on  $B(x, \delta)$  and we have*

$$(2.12) \quad \|\nabla f(x) - \nabla f(y)\| \leq \alpha\|x - y\|.$$

We finally recall basic facts from convex analysis about the Hopf-Lax-Yosida semigroup (also called Moreau envelopes).

**Definition 2.13.** Given a convex lower semicontinuous function  $g : H \rightarrow \mathbb{R}$  on a Hilbert space  $H$ , the Hopf-Lax-Yosida semigroup is the family  $g_\lambda : H \rightarrow \mathbb{R}, \lambda > 0$  defined by:

$$g_\lambda(x) = \inf_{y \in H} \frac{1}{2\lambda} \|x - y\|^2 + g(y).$$

The key result is the following:

**Proposition 2.14.** *For any continuous convex function  $g : H \rightarrow \mathbb{R}$ ,  $g_\lambda \in C^{1,1}(H)$  and is convex and for any  $x \in H$ ,  $g_\lambda(x) \rightarrow_{\lambda \rightarrow 0} g(x)$ . In fact,  $A_\lambda = \nabla g_\lambda$  is Lipschitz with constant  $\frac{1}{\lambda}$ ,  $\|A_\lambda(x)\|$  increases to  $\|A^0(x)\|$  where  $A^0(x)$  is the unique element of  $\partial g(x)$  of minimal norm and we have the inequalities:*

$$\begin{aligned} |g_\lambda(x) - g_\lambda(y) - \langle A_\lambda, x - y \rangle| &\leq \frac{1}{\lambda} \|x - y\|^2, \\ \|A_\lambda(x) - A^0(x)\|^2 &\leq \|A^0(x)\|^2 - \|A_\lambda(x)\|^2. \end{aligned}$$

Finally, for any  $x \in H$  there is a unique solution  $J_\lambda(x)$  such that  $x - J_\lambda(x) \in \lambda \partial g(J_\lambda(x))$  and  $J_\lambda : H \rightarrow H$  is a contraction such that  $A_\lambda(x) = \frac{x - J_\lambda(x)}{\lambda} \in \partial g(J_\lambda(x))$  and  $J_\lambda$  reaches the infimum defining  $g_\lambda$ .

The proof is contained in [B, Propositions 2.6, 2.11]. Note also that from the characterization of the minimum defining  $g_\lambda$ ,  $y = J_\lambda(0)$  is such that  $\frac{1}{2\lambda} \|y\|_2^2 + g(y) \leq g(0)$  so that

$$(2.13) \quad \|J_\lambda(0)\|_2 \leq 2\lambda(g(0) - g(y)).$$

We also prove the following regularity lemma in terms of parameters. It will be used to solve free SDEs with gradient drift coming from a convex potential of weak regularity via Yosida approximation in section 4.2.

**Lemma 2.15.** *If  $g^t : H \rightarrow \mathbb{R}, t \in [a, b], |b - a| \leq 1$  is a family of convex  $C^1$  maps uniformly bounded below by  $c \in \mathbb{R}$  satisfying for some  $\alpha, \beta \in ]0, 1], C, D > 0$  and all  $x, y \in H, t, s \in [a, b]$ :*

$$\begin{aligned} \|\nabla g^t(x) - \nabla g^s(x)\| &\leq |t - s|^\alpha (D\|x\| + C), \\ \|\nabla g^t(x) - \nabla g^t(y)\| &\leq \|x - y\|^\beta (D\|x\| + D\|y\| + C), \end{aligned}$$

then we have for all  $\lambda \leq 1$ :

$$\begin{aligned} \|\nabla g_\lambda^t(x) - \nabla g_\lambda^s(x)\| &\leq |t - s|^\alpha [(D\|x\|_2 + 2D(g(0) + |c|) + C) + (2D\|x\|_2 + 4D(g(0) + |c|) + C)^{1+\beta}], \\ \|\nabla g_\lambda^t(x) - \nabla g_\lambda^t(y)\| &\leq \|x - y\|^\beta (D\|x\| + D\|y\| + 4D(g(0) + |c|) + C). \end{aligned}$$

*Proof.* From the previous proposition, if we call  $J_{t,\lambda} = (1 + \lambda \partial g_t)^{-1}$ , we have since  $g_t$  is  $C^1$ :

$$\nabla g_\lambda^t(x) = \nabla g^t(J_{t,\lambda}(x))$$

Note that

$$(1 + \lambda \partial g_t)(\partial g_t)[J_{t,\lambda}(x)] = (\partial g_t)(1 + \lambda \partial g_t)[J_{t,\lambda}(x)] = (\partial g_t)(x)$$

and therefore by uniqueness of the equation characterizing  $J_{t,\lambda}$  (which comes from the uniqueness of the minimizer defining  $g_\lambda^t$ ) we have:  $(\partial g_t)[J_{t,\lambda}(x)] = J_{t,\lambda}(\partial g_t(x))$ . We can write a resolvent like equation and deduce bounds from contractivity of  $J_{t,\lambda}$ , estimate (2.13) and the assumption for  $\lambda \leq 1$ :

$$\begin{aligned} \|J_{t,\lambda}(x) - J_{s,\lambda}(x)\|_2 &= \|J_{t,\lambda}[(1 + \lambda \partial g_s)(J_{s,\lambda}(x))] - J_{t,\lambda}[(1 + \lambda \partial g_t)(J_{s,\lambda}(x))]\|_2 \\ &\leq \|\lambda(\partial g_s - \partial g_t)(J_{s,\lambda}(x))\|_2 \\ &\leq \lambda |t - s|^\alpha (D\|(J_{s,\lambda}(x))\|_2 + C) \\ &\leq \lambda |t - s|^\alpha (D\|x\|_2 + 2D(g(0) + |c|) + C). \end{aligned}$$

Combining this and the defining equation, one gets the expected result:

$$\begin{aligned} \|\nabla g_\lambda^t(x) - \nabla g_\lambda^s(x)\|_2 &\leq \|\nabla g^t(J_{t,\lambda}(x)) - \nabla g^s(J_{t,\lambda}(x))\|_2 + \|\nabla g^t(J_{t,\lambda}(x)) - \nabla g^t(J_{s,\lambda}(x))\|_2 \\ &\leq \lambda |t - s|^\alpha (D\|x\|_2 + 2D(g(0) + |c|) + C) + \|J_{s,\lambda}(x) - J_{t,\lambda}(x)\|^\beta (D\|J_{t,\lambda}(x)\| + D\|J_{s,\lambda}(x)\| + C) \\ &\leq \lambda |t - s|^\alpha (D\|x\|_2 + C) + |t - s|^{\alpha+\beta} (D\|x\|_2 + 2D(g(0) + |c|) + C)^\beta (2D\|x\|_2 + 4D(g(0) + |c|) + C) \\ &\leq \lambda |t - s|^\alpha [(D\|x\|_2 + 2D(g(0) + |c|) + C) + (2D\|x\|_2 + 4D(g(0) + |c|) + C)^{1+\beta}]. \end{aligned}$$

where the last expected conclusion is for  $\lambda \leq 1$ . Finally note that the Hölder continuity in space is obvious:

$$\begin{aligned} \|\nabla g_\lambda^t(x) - \nabla g_\lambda^t(y)\| &= \|\nabla g^t(J_{t,\lambda}(x)) - \nabla g^t(J_{t,\lambda}(y))\| \\ &\leq \|x - y\|^\beta (D\|J_{t,\lambda}(x)\| + D\|J_{t,\lambda}(y)\| + C) \\ &\leq \|x - y\|^\beta (D\|x\| + D\|y\| + 4D(g(0) + |c|) + C). \end{aligned}$$

□

**2.7. Classical Entropy.** Recall that the entropy of a probability measure  $\mu$  on  $\mathbb{R}^p$  is the quantity

$$\text{Ent}(\mu) = \begin{cases} -\int_{\mathbb{R}^p} f(x) \log f(x) dx & \text{if } \mu(dx) = f(x) dx \\ -\infty & \text{if } \mu \text{ is not absolutely continuous} \end{cases}$$

The entropy is a concave upper semi-continuous function of  $\mu$ .

Moreover, there is also a well known notion of relative entropy of two probability measures, say on a locally compact space  $\Omega$  (also called Kullback-Leibler divergence, cf. [K]).

$$\text{Ent}(\mu|\nu) = \begin{cases} -\int_{\Omega} f(x) \log f(x) d\nu(x) & \text{if } \mu(dx) = f(x) d\nu(x) \\ -\infty & \text{if } \mu \text{ is not absolutely continuous with respect to } \nu \end{cases}$$

Note that, by Jensen inequality,  $\text{Ent}(\mu|\nu) \leq 0$ .

We shall need another characterization of entropy, through its Legendre transform. Indeed one has, for any probability measure  $\mu$  supported by a set  $E$ , of finite Lebesgue measure,

$$\text{Ent}(\mu) = \inf_{\phi \in C_b(E)} \left( \log \left( \int_E \exp \phi(x) dx \right) - \int_E \phi(x) \mu(dx) \right).$$

Likewise (see e.g. [DZ, section 6.2] ) for any probability measures  $\mu, \nu$  supported on  $E$ ,

$$(2.14) \quad \text{Ent}(\mu|\nu) = \inf_{\phi \in C_b(E)} \left( \log \left( \int_E \exp \phi(x) d\nu(x) \right) - \int_E \phi(x) \mu(dx) \right).$$

We will even need a stronger representation in case of gaussian measures on  $\mathbb{R}^p$ . Let  $C_{sq}(E)$  the space of continuous functions  $\phi$  subquadratic in the sense that there is  $C$  such that  $|\phi(x)| \leq C\|x\|^2$ .

If  $\nu$  is a gaussian measure,  $\int_E \exp(C\|x\|^2) d\nu(x) < \infty$  and  $-M \vee \phi \wedge M \rightarrow \phi$  so that by dominated convergence  $\int_E \exp(-M \vee \phi(x) \wedge M) d\nu(x) \rightarrow_{M \rightarrow \infty} \int_E \exp \phi(x) d\nu(x)$ . Moreover, if  $\mu$  has a second moment by dominated convergence theorem again  $\int_E -M \vee \phi(x) \wedge M \mu(dx) \rightarrow_{M \rightarrow \infty} \int_E \phi(x) \mu(dx)$ . Let

$$\mathcal{P}^2(\mathbb{R}^p) = \{\mu \in \mathcal{P}(\mathbb{R}^p) : \int_{\mathbb{R}^p} \|x\|_2^2 d\mu(x) < \infty\}$$

Thus one deduces that for  $\nu$  gaussian measure,  $\mu \in \mathcal{P}^2(\mathbb{R}^p)$ :

$$(2.15) \quad \text{Ent}(\mu|\nu) = \inf_{\phi \in C_{sq}(E)} \left( \log \left( \int_E \exp \phi(x) d\nu(x) \right) - \int_E \phi(x) \mu(dx) \right).$$

Finally, if  $\mu$  is the restriction of  $\nu$  to  $E$ , renormalized into a probability measure, then

$$\text{Ent}(\mu|\nu) = \log(\nu(E))$$

and again this is the maximum value of  $\text{Ent}(\cdot|\nu)$  on the set of all probability measures supported by  $E$ .



**2.8. Free entropy.** Let  $\tau \in \mathcal{S}_R^{m+\mu} \star \mathcal{T}(\mathcal{F}_1^\nu)$  (cf subsection 2.5), let  $\epsilon > 0$  be a real number and  $K, N$  be positive integers. Let  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))^\nu$ . We denote by  $\Gamma_{R,\Upsilon}(\tau, \epsilon, K, N)$  the set of  $n + \mu$ -tuples of hermitian matrices  $M_1, \dots, M_{m+\mu} \in H_N^R$  such that for all monomials  $m(X_1, \dots, X_{m+\mu}, v_1, \dots, v_\nu) \in \mathbb{C}\langle X_1, \dots, X_{m+\mu}, v_1, \dots, v_\nu \rangle$  of degree less than  $K$  one has

$$|\tau(m(X_1, \dots, X_{m+\mu}, v)) - \frac{1}{N} \text{Tr}(m(M_1, \dots, M_{m+\mu}, \Upsilon))| < \epsilon$$

Equivalently  $\Gamma_{R,\Upsilon}(\tau, \epsilon, K, N)$  is the set of  $n + \mu$ -tuples of hermitian matrices  $M_1, \dots, M_{m+\mu} \in H_N^R$  whose associated state  $\sigma_{M_1, \dots, M_{m+\mu}, \Upsilon} \in \mathcal{S}_R^{m+\mu} \star \mathcal{T}(\mathcal{F}_1^\nu)$ , defined by

$$\sigma_{M_1, \dots, M_{m+\mu}, \Upsilon}(P) = \frac{1}{N} \text{Tr}(P(M_1, \dots, M_{m+\mu}, \Upsilon)),$$

is in  $V_{\epsilon, K}(\tau)$ , defined in subsection 2.7. We will write similarly as  $\sigma_{X,v}$  any mixed law of self-adjoint variables  $X$  and unitaries  $v$  from any tracial von Neumann algebra. Voiculescu defined free entropy in term of Lebesgue measure  $\text{Leb}$  on hermitian matrices, but a related definition can be made in terms of a Gaussian measure  $P$  law of  $H_1^N$  for our hermitian Brownian motion  $H_t^N$ . One can also define a third version associated to the unitary transformation we used following [BCG] :

$$u(X) = \frac{X + 4i}{X - 4i}.$$

Then the relevant set of unitaries are defined using the notation  $U_{\epsilon, K}(\sigma)$  in subsection 2.5, for  $\sigma \in \mathcal{T}_R(\mathcal{F}_1^{m+\mu}) \star \mathcal{T}(\mathcal{F}_1^\nu)$  by:

$$\begin{aligned} & \Gamma_{R,\Upsilon}^U(\sigma, \epsilon, K, N) \\ &= \{(U_1, \dots, U_m) \in U(N)^{m+\mu} : \tau_{U_1, \dots, U_{m+\mu}, \Upsilon} \in U_{\epsilon, K}(\sigma), \frac{U_j + U_j^*}{2} \leq 1 - \frac{2}{R^2 + 1}\}. \end{aligned}$$

Of course in this context we need to consider  $\Psi(X_1, \dots, X_m) = (u(X_1), \dots, u(X_m))$  and the push-forward measure  $\Psi_*P$  of our gaussian measure  $P$ . It is reasonable to give a name  $\mathcal{T}_R(\mathcal{F}_1^{m+\mu}) \star \mathcal{T}(\mathcal{F}_1^\nu)$  to the set of states  $\sigma$  such that in the GNS representation  $\frac{U_j + U_j^*}{2} \leq 1 - \frac{2}{R^2 + 1}$ . Note this is a closed set in the topology given by either  $d$  or  $d_2$ . Voiculescu introduced  $\limsup$  and  $\liminf$  variants. Moreover [S02] defined a notion of relative entropy, we define a variant where the extra variables are unitaries instead of self-adjoints (variant which is of course completely equivalent, thanks to the von Neumann algebra invariance in the relative variable, and only better suited with our framework using unitary variables). We call  $p_{m,\mu} : M_N(\mathbb{C})^{m+\nu} \rightarrow M_N(\mathbb{C})^m$  the projection on the  $m$  first coordinates.

**Definition 2.16.** [V2, S02] Let  $(M, \tau)$  a finite von Neumann algebra,  $R \in [0, \infty]$ ,  $X_1, \dots, X_m, Y_1, \dots, Y_\mu \in (M, \tau)$  self-adjoints with  $\|X_i\|, \|Y_i\| \leq R$ ,  $U_1, \dots, U_m, V_1, \dots, V_\mu, v_1, \dots, v_\nu \in \mathcal{U}(M)$ . We also fix  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^\nu$  a sequence approximating in law  $v$ . Define the various *free entropies of  $X = (X_1, \dots, X_m)$  in presence of  $Y = (Y_1, \dots, Y_\mu)$  (resp. of  $U$  in the presence of  $V$ ) relative to  $v = (v_1, \dots, v_\nu)$  (resp.  $(\Upsilon_N)$  with bound  $R$ :*

$$\begin{aligned} \chi_R(X : Y|v) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \sup_{\tau_\Upsilon \in V_{\epsilon, K}(\tau_v)} \log(\text{Leb}(p_{m,\mu} \Gamma_{R,\Upsilon}(\sigma_{X,Y,v}, \epsilon, K, N))) + \frac{m}{2} \log N \right) \\ \underline{\chi}_R(X : Y|v) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \sup_{\tau_\Upsilon \in V_{\epsilon, K}(\tau_v)} \log(\text{Leb}(p_{m,\mu} \Gamma_{R,\Upsilon}(\sigma_{X,Y,v}, \epsilon, K, N))) + \frac{m}{2} \log N \right) \\ \chi_R(X : Y|(\Upsilon_N)_{N \in \mathbb{N}}) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \log(\text{Leb}(p_{m,\mu} \Gamma_{R,\Upsilon_N}(\sigma_{X,Y,v}, \epsilon, K, N))) + \frac{m}{2} \log N \right) \\ \underline{\chi}_R(X : Y|(\Upsilon_N)_{N \in \mathbb{N}}) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \log(\text{Leb}(p_{m,\mu} \Gamma_{R,\Upsilon_N}(\sigma_{X,Y,v}, \epsilon, K, N))) + \frac{m}{2} \log N \right) \\ \chi_R^G(X : Y|v) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \sup_{\tau_\Upsilon \in V_{\epsilon, K}(\tau_v)} \log(P(p_{m,\mu} \Gamma_{R,\Upsilon}(\sigma_{X,Y,v}, \epsilon, K, N))) \right) \end{aligned}$$

$$\begin{aligned}
\underline{\chi}_R^G(X : Y|v) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \sup_{\tau_Y \in V_{\epsilon, K}(\tau_v)} \log(P(p_{m, \mu} \Gamma_{R, Y}(\sigma_{X, Y, v}, \epsilon, K, N))) \right) \\
\chi_R^G(X : Y|(\Upsilon_N)_{N \in \mathbb{N}}) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \log(P(p_{m, \mu} \Gamma_{R, Y_N}(\sigma_{X, Y, v}, \epsilon, K, N))) \right) \\
\underline{\chi}_R^G(X : Y|(\Upsilon_N)_{N \in \mathbb{N}}) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \log(P(p_{m, \mu} \Gamma_{R, Y_N}(\sigma_{X, Y, v}, \epsilon, K, N))) \right) \\
\tilde{\chi}_R(U : V|v) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \sup_{\tau_Y \in V_{\epsilon, K}(\tau_v)} \log(\Psi_* P(p_{m, \mu} \Gamma_{R, Y}^U(\tau_{U, V, v}, \epsilon, K, N))) \right) \\
\tilde{\underline{\chi}}_R(U : V|v) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \sup_{\tau_Y \in V_{\epsilon, K}(\tau_v)} \log(\Psi_* P(p_{m, \mu} \Gamma_{R, Y}^U(\tau_{U, V, v}, \epsilon, K, N))) \right) \\
\tilde{\chi}_R(U : V|(\Upsilon_N)_{N \in \mathbb{N}}) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \log(\Psi_* P(p_{m, \mu} \Gamma_{R, Y_N}^U(\tau_{U, V, v}, \epsilon, K, N))) \right) \\
\tilde{\underline{\chi}}_R(U : V|(\Upsilon_N)_{N \in \mathbb{N}}) &= \lim_{K \rightarrow \infty, \epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \log(\Psi_* P(p_{m, \mu} \Gamma_{R, Y_N}^U(\tau_{U, V, v}, \epsilon, K, N))) \right)
\end{aligned}$$

The *free entropy of  $X = (X_1, \dots, X_m)$  in presence of  $Y = (Y_1, \dots, Y_\mu)$  (resp. of  $U$  in the presence of  $V$ ) relative to  $v = (v_1, \dots, v_\nu)$  (resp.  $\Upsilon = (\Upsilon_N)_{N \in \mathbb{N}}$ , resp. a subalgebra  $B \subset M$ ) is for  $\chi \equiv \chi^L$  and  $p \in \{L, G\}$ :*

$$\begin{aligned}
\chi^p(X : Y|v) &= \sup_{R > 0} \chi_R^p(X : Y|v), \quad \tilde{\chi}(U : V|v) = \sup_{R > 0} \tilde{\chi}_R(U : V|v), \\
\underline{\chi}^p(X : Y|v) &= \sup_{R > 0} \underline{\chi}_R^p(X : Y|v), \quad \tilde{\underline{\chi}}(U : V|v) = \sup_{R > 0} \tilde{\underline{\chi}}_R(U : V|v),
\end{aligned}$$

$$\begin{aligned}
\chi^p(X : Y|\Upsilon) &= \sup_{R > 0} \chi_R^p(X : Y|(\Upsilon_N)_{N \in \mathbb{N}}), \quad \tilde{\chi}(U : V|\Upsilon) = \sup_{R > 0} \tilde{\chi}_R(U : V|(\Upsilon_N)_{N \in \mathbb{N}}), \\
\underline{\chi}^p(X : Y|\Upsilon) &= \sup_{R > 0} \underline{\chi}_R^p(X : Y|(\Upsilon_N)_{N \in \mathbb{N}}), \quad \tilde{\underline{\chi}}(U : V|\Upsilon) = \sup_{R > 0} \tilde{\underline{\chi}}_R(U : V|(\Upsilon_N)_{N \in \mathbb{N}}).
\end{aligned}$$

$$\begin{aligned}
\chi^p(X : Y|B) &= \inf_{\nu \in \mathbb{N}} \inf_{v_1, \dots, v_\nu \in \mathcal{U}(B)} \chi^p(X : Y|v), \quad \tilde{\chi}(U : V|v) = \inf_{\nu \in \mathbb{N}} \inf_{v_1, \dots, v_\nu \in \mathcal{U}(B)} \tilde{\chi}(U : V|v), \\
\underline{\chi}^p(X : Y|B) &= \inf_{\nu \in \mathbb{N}} \inf_{v_1, \dots, v_\nu \in \mathcal{U}(B)} \underline{\chi}^p(X : Y|v), \quad \tilde{\underline{\chi}}(U : V|B) = \inf_{\nu \in \mathbb{N}} \inf_{v_1, \dots, v_\nu \in \mathcal{U}(B)} \tilde{\underline{\chi}}(U : V|v).
\end{aligned}$$

Of course we will write  $\chi(X|\cdot)$  if  $\mu = 0$  and  $\chi(X : Y)$  if  $\nu = 0$  and similar variants.

As is well-known and as was noticed e.g. in [BCG, section 7], the above entropies are related by a universal constant  $C$  such that:

$$(2.16) \quad \chi(X : Y|\cdot) = \chi^G(X : Y|\cdot) + \frac{1}{2} \sum_{i=1}^m \tau(X_i^2) + mC.$$

and their lemma 7.1 shows that if  $\Psi(X_1, \dots, X_m) = (u(X_1), \dots, u(X_m))$

$$(2.17) \quad \chi^G(X : Y|\cdot) = \tilde{\chi}(\Psi(X) : \Psi(Y)|\cdot).$$

The analogue formulas for  $\liminf$  variants and for  $\chi_\infty$  variants are also true.

As in [S02, Theorem 2.15] (a consequence of Kaplansky density theorem),

$$(2.18) \quad \chi(X : Y|u_1, \dots, u_\nu) = \chi(X : Y|W^*(u_1, \dots, u_\nu)),$$

and this last version is the same as the variant defined using self-adjoint variables in this paper. This algebra invariance can also be deduced from Theorem E for which the proof will be more detailed. All the other variants are also true for instance with  $\chi$  replaced by  $\chi^G$ .

Finally, among known results if  $v$  generates a hyperfinite algebra and  $\Upsilon = (\Upsilon_N)_{N \in \mathbb{N}}$  has a law tending to  $v$ , then (see e.g. the proof of [Ue14, Prop 2.3, 2.4] inspired from [HMU, lemma 4.2]), then  $\chi(X : Y|v) = \chi(X : Y|\Upsilon)$ . We will generalize this beyond the hyperfinite case when  $\{Y\} = \emptyset$ .

The following result from [BB] will be crucial to apply large deviation principle to free entropy. The proof gives right away the result for  $\chi, \chi^G$  and their lim inf variants, and then (2.17) deals with the remaining case.

**Proposition 2.17** (Proposition 2.1 in [BB]). *We have in the setting of the previous definition,  $p \in \{L, G\}$ :*

$$\begin{aligned}\chi^p(X : Y|\cdot) &= \chi_\infty^p(X : Y|\cdot), \quad \tilde{\chi}(\Psi(X) : \Psi(Y)|\cdot) = \tilde{\chi}_\infty(\Psi(X) : \Psi(Y)|\cdot), \\ \underline{\chi}^p(X : Y|\cdot) &= \underline{\chi}_\infty^p(X : Y|\cdot), \quad \underline{\tilde{\chi}}(\Psi(X) : \Psi(Y)|\cdot) = \underline{\tilde{\chi}}_\infty(\Psi(X) : \Psi(Y)|\cdot).\end{aligned}$$

We will also need the following special case of [S02, Th 2.19]. Since the proof there is not convincing (in absence of our proof that the lim sup and lim inf versions of free entropy coincide), we include the standard proof.

**Proposition 2.18.** *Let  $X_1, X_2$ , selfadjoints and  $v$  unitaries in  $(M, \tau)$  as above, then if  $\{X_1, v\}$  and  $X_2$  are free we have :*

$$\begin{aligned}\chi(X_1, X_2|v) &= \chi(X_1|v) + \chi(X_2), \\ \underline{\chi}(X_1, X_2|v) &= \underline{\chi}(X_1|v) + \chi(X_2).\end{aligned}$$

*Proof.* In both cases,  $\leq$  are general (see [S02, Prop 2.18] for the lim sup case). Since lim inf is not subadditive but superadditive, one uses that the lim sup and the lim inf variants of entropy are equal in the case  $n = 1$  (see e.g. [HP, Th 5.6.2]). From those inequalities, we can assume  $\chi(X_1, v) > -\infty, \chi(X_2) > -\infty$ . Fix  $R > \|X_1\|, \|X_2\|$ ,  $\Upsilon$  with  $\tau_\Upsilon \in V_{\epsilon, K}(\tau_v)$ .

We let

$$\Psi(N, K, \epsilon) = \Gamma_{R, \Upsilon}(\sigma_{X_1, X_2, v}, \epsilon, K, N), \quad \Phi(N, K, \epsilon) = \Gamma_{R, \Upsilon}(\sigma_{X_1, v}, \epsilon, K, N) \times \Gamma_{R, \Upsilon}(\sigma_{X_2}, \epsilon, K, N).$$

As in [V1, lemma 3.5] for each  $K \in \mathbb{N}, \epsilon > 0$  there is  $\epsilon > \delta_0 > 0$  such that for all sufficiently large  $N$   $\delta < \delta_0$ , whatever  $\Upsilon$  chosen as above with also  $\tau_\Upsilon \in V_{\delta, K}(\tau_v)$ :

$$\frac{\text{Leb}(\Psi(N, K, \epsilon) \cap \Phi(N, K, \delta))}{\text{Leb}(\Phi(N, K, \delta))} \geq \frac{1}{2}.$$

Thus one gets in taking a log and a supremum:

$$\begin{aligned}\sup_{\tau_\Upsilon \in V_{\epsilon, K}(\tau_v)} \frac{1}{N^2} \log(\text{Leb}(\Gamma_{R, \Upsilon}(\sigma_{X_1, X_2, v}, \epsilon, K, N))) &\geq \sup_{\tau_\Upsilon \in V_{\delta, K}(\tau_v)} \frac{1}{N^2} \log(\text{Leb}(\Psi(N, K, \epsilon) \cap \Phi(N, K, \delta))) \\ &\geq \sup_{\tau_\Upsilon \in V_{\delta, K}(\tau_v)} \frac{1}{N^2} \log(\text{Leb}(\Gamma_{R, \Upsilon}(\sigma_{X_1, v}, \delta, K, N))) + \frac{1}{N^2} \log(\Gamma_{R, \Upsilon}(\sigma_{X_2}, \delta, K, N)) - \frac{1}{N^2} \log(2).\end{aligned}$$

Adding  $\log(N)$ , taking a lim sup or a lim inf in  $N$  (in putting the inequalities in the right form to use respectively subadditivity and superadditivity) and then an infimum (which are limits) over  $\delta$ , then  $\epsilon, K$  and finally a supremum over  $R$ , one gets:

$$\begin{aligned}\chi(X_1, X_2|v) &\geq \chi(X_1|v) + \underline{\chi}(X_2), \\ \underline{\chi}(X_1, X_2|v) &\geq \underline{\chi}(X_1|v) + \underline{\chi}(X_2).\end{aligned}$$

Since  $\underline{\chi}(X_2) = \chi(X_2)$ , one concludes.  $\square$

In order to recall the definition of Voiculescu's non-microstate free entropy, we recall first the definition of free Brownian motion. Here  $Id_{\mathbb{R}^m}$  is the unit in  $M_n(\mathbb{R})$

**Definition 2.19.** Let  $B_s$  be an increasing filtration of von Neumann algebras in a non-commutative tracial probability space  $(M, \tau)$ .  $S_s = (S_s^1, \dots, S_s^m), s \in \mathbb{R}_+$  an  $m$ -tuple of self-adjoint processes adapted to this filtration with  $Z_0 = 0$  is a *free brownian motion* adapted for  $B_s$  if:

- (1)  $(S_t - S_s)$  are free semi-circular variables of covariance  $(t - s)Id_{\mathbb{R}^m}$ .
- (2)  $\{(S_u - S_s), u \geq s\}$  are free from  $B_s$ .

There is an important characterization of free brownian motion in the spirit of Paul Lévy's characterization of ordinary brownian motion. It is due to [BCG].

**Theorem 2.20** (Theorem 6.2 in [BCG]). *Let  $B_s$  be an increasing filtration of von Neumann algebras in a non-commutative tracial probability space  $(M, \tau)$ .  $Z_s = (Z_s^1, \dots, Z_s^m)$ ,  $s \in \mathbb{R}_+$  an  $m$ -tuple of self-adjoint processes adapted to this filtration with  $Z_0 = 0$  and :*

- (1)  $\tau(Z_t|B_s) = Z_s$
- (2)  $\tau(|Z_t - Z_s|^4) \leq K(t-s)^2$  for some constant  $K > 0$ .
- (3)  $\tau(Z_t^k A Z_t^l B) = \tau(Z_s^k A Z_s^l B) + (t-s)\tau(A)\tau(B)1_{\{k=l\}} + o(t-s)$  for any  $A, B \in B_s$ .

Then  $Z$  is a free brownian motion adapted to  $B_s$ .

Let us recall the definition of free entropy relative to a subalgebra  $B$  from [V5]. Here  $S_1, \dots, S_m$  are free semicircular variables free from  $X_1, \dots, X_m, B$ :

$$\chi^*(X_1, \dots, X_m : B) = \frac{1}{2} \int_0^\infty \left( \frac{m}{1+t} - \Phi^*(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m : B) \right) dt + \frac{m}{2} \log(2\pi e).$$

Let us also remind the Fisher information  $\Phi^*(Y_1, \dots, Y_m : B) = \sum_{i=1}^m \|\xi_i\|_2^2$  where  $\xi_i$  are the conjugate variables relative to  $B$  (which exists for  $Y = X + \sqrt{t}S$  as above, and are (when they exist) the unique  $\xi_i \in L^2(W^*(B, Y_1, \dots, Y_m))$  such that for all  $P \in B\langle X_1, \dots, X_m \rangle$ , if  $\partial_i$  is the free difference quotient, unique derivation with  $\partial_i(b) = 0, b \in B$  and  $\partial_i X_j = 1 \otimes 1_{i=j} \in L^2(W^*(B, Y_1, \dots, Y_m) \otimes W^*(B, Y_1, \dots, Y_m))$ , then:

$$\langle 1 \otimes 1, \partial_i P \rangle = \tau(\xi_i P).$$

The gaussian variant is defined using the gaussian variant of Fisher's information,  $\Phi^{G*}(Y_1, \dots, Y_m : B) = \sum_{i=1}^m \|\xi_i - Y_i\|_2^2$  by:

$$\chi^{G*}(X_1, \dots, X_m : B) = -\frac{1}{2} \int_0^1 \frac{dt}{t} \Phi^{G*}(\sqrt{t}X_1 + \sqrt{1-t}S_1, \dots, \sqrt{t}X_m + \sqrt{1-t}S_m : B).$$

It is easy to see in using linear changes of variables for the score function from [V5] that  $\chi^{G*}, \chi^*$  are related by (2.16).

**2.9. Ultraproducts.** In this final preliminary subsection, we recall backgrounds on ultraproducts. We refer to [P] for more details in the tracial von Neumann algebra context and to [Fa, CL] in the model theory context. Let  $(M_n, \tau_n)$  a sequence of tracial von Neumann algebras. We will mainly use the case  $M_n = M_n(L^\infty(\Omega_n, P_n))$  of matrix algebras over a classical probability space  $(\Omega_n, P_n)$ , with  $\tau_n = E \circ \frac{1}{n} \text{Tr}_n$ . Let  $\omega \in \beta\mathbb{N} - \mathbb{N}$  a non-principal ultrafilter (or equivalently a non-integer point in the Stone-Cech compactification  $\beta\mathbb{N}$  of  $\mathbb{N}$ ).

The ultraproduct of this sequence is defined as the following quotient of the set of bounded sequences with the  $n$ -th term of the sequence in  $M_n$ , noted  $\ell^\infty(M_n, n \in \mathbb{N})$ :

$$(M_n, \tau_n)^\omega = \ell^\infty(M_n, n \in \mathbb{N}) / \{(x_n) : \lim_{n \rightarrow \omega} \tau_n(x_n^* x_n) = 0\}.$$

It is known that  $(M_n, \tau_n)^\omega$  is a tracial von Neumann algebra with trace :

$$\tau_\omega((x_n)^\omega) = \lim_{n \rightarrow \omega} \tau_n(x_n).$$

There is a recent model theoretic proof of this fact [Fa], but the classical proof (see e.g. [P, section 9.10]) gives an explicit action on a Hilbert space that we will need. Actually, in writing  $(M_n, \tau_n)^\omega$  we consider always given the trace  $\tau_\omega$ .

Let  $H_n = L^2(M_n, \tau_n)$  the Hilbert space of the GNS representation. For instance, if  $M_n = M_n(L^\infty(\Omega_n, P_n))$ ,  $L^2(M_n, \tau_n) = M_n(L^2(\Omega_n, P_n))$  with its canonical Hilbert space structure with scalar product  $\langle u, v \rangle_2 = E(\frac{1}{n} \text{Tr}(u^* v))$ .

Consider the Hilbert space ultraproduct :

$$(L^2(M_n, \tau_n))^\omega = \ell^\infty(L^2(M_n, \tau_n); n \in \mathbb{N}) / \{(h_n), \lim_{n \rightarrow \omega} \|h_n\| \rightarrow 0\}$$

Then  $(M_n, \tau_n)^\omega \subset B((L^2(M_n, \tau_n))^\omega)$  with the action given by

$$(x_n)^\omega (h_n)^\omega = (x_n h_n)^\omega.$$

Then the GNS construction  $L^2((M_n, \tau_n)^\omega)$  is a subset of  $(L^2(M_n, \tau_n))^\omega$  that can be described as follows (see e.g. [P, Rmk 9.10.2]). This is either the closure of  $(M_n, \tau_n)^\omega$  (included by its action

on  $(1)^\omega$ ). Alternatively, a sequence  $(h_n)^\omega \in (L^2(M_n, \tau_n))^\omega$  belongs to  $L^2((M_n, \tau_n)^\omega)$  if and only if the following uniform integrability like condition holds :

$$\lim_{c \rightarrow \infty} \lim_{n \rightarrow \omega} \tau_n(h_n^* h_n 1_{h_n^* h_n \geq c}) = 0.$$

Here  $1_{h \geq c}$  denotes the spectral projection of the positive operator  $h \in L^1(M_n, \tau_n)$ .

### 3. APPLICATION OF THE BOUÉ-DUPUIS-ÜSTÜNEL FORMULA

In [BD], Boué and Dupuis proved a formula for exponential functionals of brownian motion. They deduced from it large deviation results, and we will use an improvement with exactly the same goal. Recall that a process on the Wiener space  $\Omega = \mathbb{W}$  is said progressively measurable if its restriction to  $[0, t] \times \Omega$  is (jointly) measurable with respect to the canonical brownian filtration  $\mathcal{F}_t$  (tensor borel sets on  $[0, t]$ ). We refer to [BD] for the following result (and also [L] for an enlightening explanation).

**Theorem 3.1.** *For every function  $f: \mathbb{W} \rightarrow \mathbb{R}$  measurable and bounded from above, we have*

$$-\log \left( \int_{\mathbb{W}} e^{-f} d\gamma \right) = \inf_{U \in L_a^2(\gamma, \mathbb{H})} \left[ \mathbf{E}_\gamma \left( f(B + U) + \frac{1}{2} \|U\|_{\mathbb{H}}^2 \right) \right],$$

where the supremum is taken over  $L_a^2(\gamma, \mathbb{H})$  of all progressively measurable processes  $U$  which belongs to  $\mathbb{H}$  almost surely.

The boundedness assumption will be annoying for our purposes since the typical convex functions we considered on matrix hermitian brownian motion  $\mathcal{E}_{reg,p}(\mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m), d_2)$  are only subquadratic. Fortunately, Üstünel extended recently this formula to a wider class of functionals [Us14]. He studied the class of functionals satisfying this theorem under the name “tame functionals”.

**Definition 3.2.** A measurable map  $f: W \rightarrow \mathbb{R} \cup \{\infty\}$ , with the property  $E_\gamma[(1 + |f|)e^{-f}] < \infty$ , is called a **tamed functional** if

$$-\log \left( \int_{\mathbb{W}} e^{-f} d\gamma \right) = \inf_{U \in L_a^2(\gamma, \mathbb{H})} \left[ \mathbf{E}_\gamma \left( [f(B + U) + \frac{1}{2} \|U\|_{\mathbb{H}}^2] \right) \right]$$

The result we need is:

**Theorem 3.3** (Theorem 7 in [Us14]). *Every measurable function  $f: \mathbb{W} \rightarrow \mathbb{R}$  such that  $f \in L^p(\gamma)$  and  $e^{-f} \in L^q(\gamma)$  with  $\frac{1}{p} + \frac{1}{q} = 1$  is a tame functional.*

This result can straightforwardly be applied to hermitian brownian motions and functionals  $W \mapsto f(\tau_X)$  with  $f \in \mathcal{E}_m(\mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m), d_2)$ . We will develop in the next section a more explicit formulation in this case, better suited for ultraproducts techniques.

We will also need the following consequence when we apply it to a brownian motion on  $[t, 1]$  for  $t \in ]0, 1[$ . Let  $\mathbb{W}_{[0,t]}, \mathbb{W}_{[t,1]}$  the spaces of continuous functions starting at zero and  $\gamma_{[0,t]}, \gamma_{[t,1]}$  the standard Wiener measure on them (for the second, we write  $(B - B_t)$  the brownian variable for consistency). For  $\omega \in \mathbb{W}_{[0,t]}, \nu \in \mathbb{W}_{[t,1]}$ , there is a process  $\omega + \nu \in \mathbb{W}$  such that  $(\omega + \nu)_s = \omega_s$  for  $s \leq t$  and  $(\omega + \nu)_s = \omega_t + \nu_s$  for  $s \geq t$ . For  $f: \mathbb{W} \rightarrow \mathbb{R}$  measurable,  $\nu \mapsto f(\omega + \nu)$  defines a measurable function.

**Corollary 3.4.** *Fix  $t \in ]0, 1[$  and a measurable function  $f: \mathbb{W} \rightarrow \mathbb{R}$  such that  $f \in L^p(\gamma)$  and  $e^{-f} \in L^q(\gamma)$  with  $\frac{1}{p} + \frac{1}{q} = 1$ , then for  $\gamma_{[0,t]}$  almost all  $\omega \in \mathbb{W}_{[0,t]}$  we have the equality :*

$$\begin{aligned} \lambda_t(f)(\omega) &:= -\log \left( \int_{\mathbb{W}_{[t,1]}} e^{-f(\omega + \nu)} d\gamma_{[t,1]}(\nu) \right) \\ (3.1) \quad &= \inf_{U \in L_a^2(\gamma_{[t,1]}, \mathbb{H})} \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [f(\omega + (B - B_t) + U) + \frac{1}{2} \|U\|_{\mathbb{H}}^2] \right) \right]. \end{aligned}$$

Moreover the condition holds for every  $\omega \in \mathbb{W}_{[0,t]}$  if  $e^{-f(\omega + \cdot)} \in L^q(\gamma_{[t,1]})$  and  $f(\omega + \cdot) \in L^p(\gamma_{[t,1]})$  for every  $\omega \in \mathbb{W}_{[0,t]}$ . As a consequence, in this case, if  $f$  is moreover convex, so is  $\lambda_t(f)$ .



*Proof.* By Fubini theorem, since  $\mathbf{E}_\gamma(e^{-qf}) = \int d\gamma_{[0,t]}(\omega) \int d\gamma_{[t,1]}(\nu) e^{-qf(\omega+\nu)}$ , we have for  $\gamma_{[0,t]}$  almost all  $\omega$ ,  $e^{-f(\omega+\cdot)} \in L^q(\gamma_{[t,1]})$  and similarly  $f(\omega+\cdot) \in L^p(\gamma_{[t,1]})$ ; The formula is thus deduced from Üstünel's theorem. For the convexity result, only take  $U_1, U_2 \in L_a^2(\gamma_{[t,1]}, \mathbb{H})$ , and note that by convexity of  $f$  and  $|\cdot|_H^2$ , one gets for  $\lambda \in [0, 1]$ :

$$\begin{aligned} & \mathbf{E}_{\gamma_{[t,1]}} \left( [f(\lambda\omega_1 + (1-\lambda)\omega_2 + (B - B_t) + \lambda U_1 + (1-\lambda)U_2) + \frac{1}{2}\|\lambda U_1 + (1-\lambda)U_2\|_{\mathbb{H}}^2] \right) \\ & \leq \lambda \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [f(\omega_1 + (B - B_t) + U_1) + \frac{1}{2}\|U_1\|_{\mathbb{H}}^2] \right) \right] \\ & \quad + (1-\lambda) \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [f(\omega_2 + (B - B_t) + U_2) + \frac{1}{2}\|U_2\|_{\mathbb{H}}^2] \right) \right]. \end{aligned}$$

This concludes to the convexity once taken various infima.  $\square$

The use of  $\lambda_s$  is standard in optimal control (cf. e.g. [FS]) to solve the minimization problem in Boué-Dupuis-Üstünel formula. It has a (non-linear) semigroup property in the form  $\lambda_s(\lambda_{s+t}(f)) = \lambda_s(f)$ , for  $s, t > 0$ ,  $s + t < 1$ . We gather this and several basic properties in our next result.

We first fix some preliminary notation. Let  $f \in C^0(\mathbb{R}^{dk}, \mathbb{R})$  a continuous function and  $t_1 < \dots < t_k \in [0, 1]$ . Let  $J_{t_1, \dots, t_k} : \mathbb{W} \rightarrow \mathbb{R}^{dk}$  such that  $J_{t_1, \dots, t_k}(\omega) = (\omega_{t_1}, \dots, \omega_{t_k})$ . We let  $\mathcal{E}(\mathbb{R}^{dk})$  the set of convex continuous functions  $g$ , bounded from below and such that

$$(3.2) \quad g(x_1, \dots, x_k) \leq c(d + \|x\|^2),$$

(as usual we wrote  $\|x\|^2 = \sum_{i=1}^k \sum_{j=1}^d (|x_i|^{(j)})^2$  the euclidean norm) and with the following local Lipschitz condition for the euclidean norm on  $\mathbb{R}^{dk}$ , with some  $C \geq 1, D \geq 0$  such that for all  $x, y$ :

$$(3.3) \quad g(x) \leq g(y) + (C\|y\| + C\|x\| + D\sqrt{d})\|x - y\|$$

For  $\alpha \in [1, 2]$ , we call  $\mathcal{E}_\alpha(\mathbb{R}^{dk})$  the subset of  $\mathcal{E}(\mathbb{R}^{dk})$  such that there exists  $C_\alpha, D_\alpha > 0$  with for all  $x, y \in \mathbb{R}^{dk}$ :

$$(3.4) \quad g(x+y) + g(x-y) - 2g(x) \leq d^{1-\alpha/2} (C_\alpha + D_\alpha \frac{\|x\| + \|x+y\| + \|x-y\|}{\sqrt{d}}) \|y\|^\alpha.$$

Note that by convexity, the left hand side is positive. The fact that one can only obtain one sided bounds for second order difference quotients of value functions is standard in optimal control (see e.g. [FS, section IV.9]). By [N75], this class of functions are Gâteaux-differentiable if  $\alpha > 1$  so that we can write  $D_H^y g(x_1, \dots, x_{n-1}, y) = \sum_{i=1}^d \frac{\partial}{\partial y^{(i)}} g(x_1, \dots, x_{n-1}, y) H_i$  for  $H \in \mathbb{R}^d$ . Note also we made appear the dimension  $d$  explicitly since we will use later families of models where the constants involved will be dimension independent once used the conventions above.

**Proposition 3.5.** *Let  $t_0 = 0 < t_1 < \dots < t_k \in [0, 1]$ ,  $g \in \mathcal{E}(\mathbb{R}^{dk})$  and  $f = g \circ J_{t_1, \dots, t_k} : \mathbb{W} \rightarrow \mathbb{R}$ , then (3.1) holds for every  $t \in [0, 1]$  and  $\omega \in \mathbb{W}_{[0,t]}$ . Moreover, if  $t_i < t \leq t_{i+1}$ , then there is  $h_t \in \mathcal{E}(\mathbb{R}^{d(i+1)})$  with  $\lambda_t(f) = h_t \circ J_{t_1, \dots, t_i, t}$ , and for all  $s, t > 0, s + t < 1$ :*

$$\lambda_s(\lambda_{s+t}(f)) = \lambda_s(f).$$

Moreover, there are constants  $c_1, d_1 > 0$  such that for all  $t, t+s \in ]t_i, t_{i+1}]$ ,  $s > 0$  we have :

$$(3.5) \quad |h_t(x_1, \dots, x_i, x) - h_{t+s}(x_1, \dots, x_i, x)| \leq \sqrt{s}(c_1\|(x_1, \dots, x_i, x)\|^2 + d_1(d + \sup(-g)))$$

and for  $t = t_i, s > 0, t+s \in ]t_i, t_{i+1}]$ ,

$$\begin{aligned} & |h_t(x_1, \dots, x_i) - h_{t+s}(x_1, \dots, x_i, x)| \\ & \leq \sqrt{(c_1\|(x_1, \dots, x_i, x)\|^2 + d_1(d + \sup(-g)))} \left( \sqrt{s(c_1\|(x_1, \dots, x_i, x)\|^2 + d_1(d + \sup(-g)))} + \|x_i - x\| \right) \end{aligned}$$

Finally, for  $\alpha \in [1, 2]$ , if  $g \in \mathcal{E}_\alpha(\mathbb{R}^{dk})$  so is  $h_t \in \mathcal{E}_\alpha(\mathbb{R}^{d(i+1)})$ , and if  $\alpha = 2$ ,  $h_t$  is Gâteaux-differentiable and if  $D_2 = 0$  for  $g$  we can take  $D_2 = 0$  for  $h_t$ , and for all  $\|H\| = 1, H \in \mathbb{R}^d$ , the following bound on the directional derivative holds:

$$|D_H^y h_t(x_1, \dots, x_i, y) - D_H^y h_{t+s}(x_1, \dots, x_i, y)| \leq s^{1/2} \sqrt{d} \left( c_1 + d_1 \frac{\|(x_1, \dots, x_i, y)\|^2 + \sup(-g)}{d} \right).$$

**Proof. Step 1 :** Formula and Stability of subquadratic behaviour

First we fix  $t$  and write  $h = h_t$ . The assumptions from the previous corollary to check (3.1) hold from the boundedness and at most quadratic growth assumption. It suffices to define:

$$h(x_1, \dots, x_i, x) = -\log \left( \int_{\mathbb{W}_{[t,1]}} e^{-g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} d\gamma_{[t,1]}(\nu) \right).$$

The expected formula  $\lambda_t(f) = h \circ J_{t_1, \dots, t_i, t}$  follows by an examination of our notation. The convexity then follows from theorem 3.3 as in the previous corollary. This results also yields the alternative formula :

$$(3.6) \quad h_t(x_1, \dots, x_i, x) = \inf_{U \in L_a^2(\gamma_{[t,1]}, \mathbb{H})} \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( g(x_1, \dots, x_i, x + [(B - B_t) + U]_{t_{i+1}}, \dots, x + [(B - B_t) + U]_{t_k}) + \frac{1}{2} \|U\|_{\mathbb{H}}^2 \right) \right].$$

Obviously,  $h = h_t$  is bounded from below with the same constant as  $g$  and taking  $U = 0$ ,

$$\begin{aligned} h(x_1, \dots, x_i, x) &\leq c \left( d + \sum_{l=1}^i \sum_{j=1}^d (|x_l|^{(j)})^2 + \mathbf{E}_{\gamma_{[t,1]}} \left( \sum_{l=i+1}^k \sum_{j=1}^d (|x + (B - B_t)_l|^{(j)})^2 \right) \right) \\ &\leq c \left( d + \sum_{l=1}^i \sum_{j=1}^d (|x_l|^{(j)})^2 + 2(k-i) \left( d + \sum_{j=1}^d |x^{(j)}|^2 \right) \right) \\ &\leq c(1 + 2(k-i)) (d + \|(x_1, \dots, x_i, x)\|^2). \end{aligned}$$

Thus  $h$  is subquadratic too.

**Step 2 :** Stability of lipschitzness

We now check the stability of (3.3). Fix  $U \in L_a^2(\gamma_{[t,1]}, \mathbb{H})$  reaching the inf up to  $\epsilon > 0$  for  $h(y_1, \dots, y_i, y)$  and use (3.3) to get as before for  $X = (x_1, \dots, x_i, x, \dots, x), Y = (y_1, \dots, y_i, y, \dots, y)$  that :

$$\begin{aligned} &\mathbf{E}_{\gamma_{[t,1]}} (g(x_1, \dots, x_i, x + [(B - B_t) + U]_{t_{i+1}}, \dots, x + [(B - B_t) + U]_{t_k})) \\ &\leq \mathbf{E}_{\gamma_{[t,1]}} (g(y_1, \dots, y_i, y + [(B - B_t) + U]_{t_{i+1}}, \dots, y + [(B - B_t) + U]_{t_k})) + \sqrt{k-i} \|x - y\| \\ &\times \mathbf{E}_{\gamma_{[t,1]}} (D\sqrt{d} + C\|X\| + 2C\|[(B - B_t) + U]_{t_{i+1}}, \dots, [(B - B_t) + U]_{t_k}\| + C\|Y\|) \\ &\leq \mathbf{E}_{\gamma_{[t,1]}} (g(y_1, \dots, y_i, y + [(B - B_t) + U]_{t_{i+1}}, \dots, y + [(B - B_t) + U]_{t_k})) + \sqrt{k-i} \|x - y\| \\ &\times \mathbf{E}_{\gamma_{[t,1]}} (D\sqrt{d} + C\|X\| + 2C\sqrt{k-i}(\sqrt{d} + \|U\|_{\mathbb{H}}) + C\|Y\|). \end{aligned}$$

We then note that we can use the definition as an infimum and the subquadratic bound above to show that a  $U$ ,  $\epsilon$  close to the infimum as we chose, satisfies:

$$\begin{aligned} (3.7) \quad \mathbf{E}(\|U\|_{\mathbb{H}}^2) &\leq 2\epsilon + 2h(y_1, \dots, y_i, y) + 2\sup(-g) \\ &\leq 2\epsilon + 2c(1 + 2(k-i)) (d + \|(y_1, \dots, y_i, y)\|^2) + 2\sup(-g) =: 2\epsilon + A(y_1, \dots, y_i, y) \end{aligned}$$

Considering such an  $U$ , we can thus combine our previous equations and after adding  $\frac{1}{2}\|U\|_{\mathbb{H}}^2$ , and taking the infemum on the left hand side, one gets:

$$h(x_1, \dots, x_i, x) \leq h(y_1, \dots, y_i, y) + \sqrt{k-i}\|x-y\| \left( D\sqrt{d} + C\|X\| + C\|Y\| \right) + \epsilon \\ + 2C(k-i)\|x-y\| \left( \sqrt{d} + \sqrt{2\epsilon + A(y_1, \dots, y_i, y)} \right)$$

thus letting  $\epsilon \rightarrow 0$  one gets (3.3) for  $h$  with  $C$  replaced by  $C_{k,i} = C(k-i) + 2C(k-i)\sqrt{2c(1+2(k-i))}$  and  $D$  replaced by  $D_{k,i} = D\sqrt{k-i} + 2C(k-i)(1 + \sqrt{2c(1+2(k-i))} + \sqrt{\frac{2}{d}\sup(-g)})$ .

**Step 3 :** “Semigroup” formula

In decomposing Wiener measure as a product measure by independence and using the defining formula twice and Fubini theorem :

$$\lambda_s(\lambda_{s+t}(f)) = -\log \left( \int_{\mathbb{W}_{[s,t+s]}} e^{-\lambda_{s+t}(f)(\omega+\nu)} d\gamma_{[s,t+s]}(\nu) \right) \\ = -\log \left( \int_{\mathbb{W}_{[s,t+s]}} \int_{\mathbb{W}_{[t+s,1]}} e^{-f(\omega+\nu+\mu)} d\gamma_{[t+s+1]}(\mu) d\gamma_{[s,t+s]}(\nu) \right) \\ = \lambda_s(f).$$

This will be the key to all our time estimates bellow.

**Step 4 :** Second order bound

Finally, let us prove the statement for  $g \in \mathcal{E}_\alpha(\mathbb{R}^{dk})$ . Consider a typical element in the infemum formula defining  $h$  for  $x, Y$  with  $Y = (y_1, \dots, y_i, y), Y' = (y_1, \dots, y_i, y, \dots, y), X' = (x_1, \dots, x_i, x, \dots, x)$  and apply (3.4) to  $g$  in order to get :

$$\left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [2g(x_1, \dots, x_i, x + [(B-B_t) + U]_{t_{i+1}}, \dots, x + [(B-B_t) + U]_{t_k}) + \|U\|_{\mathbb{H}}^2] \right) \right] \\ \geq \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [g(x_1 + y_1, \dots, x_i + y_i, x + y + [(B-B_t) + U]_{t_{i+1}}, \dots, x + y + [(B-B_t) + U]_{t_k}) + \frac{1}{2}\|U\|_{\mathbb{H}}^2] \right) \right] \\ + \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [g(x_1 - y_1, \dots, x_i - y_i, x - y + [(B-B_t) + U]_{t_{i+1}}, \dots, x - y + [(B-B_t) + U]_{t_k}) + \frac{1}{2}\|U\|_{\mathbb{H}}^2] \right) \right] \\ - \left[ d^{1-\alpha/2} (D_\alpha \frac{\|X'\| + \|X' + Y'\| + \|X' - Y'\|}{\sqrt{d}}) \|Y'\|^\alpha \right] \\ - \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( d^{1-\alpha/2} (C_\alpha + D_\alpha \frac{3\|[(B-B_t) + U]_{t_{i+1}}, \dots, x + [(B-B_t) + U]_{t_k}\|}{\sqrt{d}}) \|Y'\|^\alpha \right) \right] \\ \geq h(x_1 + y_1, \dots, x_i + y_i, x + y) + h(x_1 - y_1, \dots, x_i - y_i, x - y) \\ - \left[ d^{1-\alpha/2} (D_\alpha \frac{\|X'\| + \|X' + Y'\| + \|X' - Y'\|}{\sqrt{d}}) \|Y'\|^\alpha \right] \\ - \left[ d^{1-\alpha/2} (C_\alpha + D_\alpha \frac{3(\sqrt{d}(k-i) + (k-i)\sqrt{\mathbf{E}_{\gamma_{[t,1]}}(\|U\|_{\mathbb{H}}^2)})}{\sqrt{d}}) \|Y'\|^\alpha \right].$$

But  $\|Y'\|^\alpha \leq (k-i)^{\alpha/2}\|Y\|^\alpha$  and one can bound from (3.7)  $\mathbf{E}_{\gamma_{[t,1]}}(\|U\|_{\mathbb{H}}^2) \leq 2\epsilon + A(x_1, \dots, x_i, x)$  so that taking the infemum,  $h$  satisfies (3.4) with  $C_\alpha$  replaced by

$$C_{\alpha,k,i} := \left( C_\alpha + 3D_\alpha(k-i) + 3D_\alpha(k-i)\sqrt{2c(1+2(k-i))} + \frac{3(k-i)D_\alpha\sqrt{2\sup(-g)}}{\sqrt{d}} \right) (k-i)^{\alpha/2},$$

and  $D_\alpha$  replaced by (a value which is 0 if  $D_\alpha = 0$ ):

$$D_{\alpha,k,i} := \left( 2D_\alpha \sqrt{k-i} + 3D_\alpha(k-i)\sqrt{2c(1+2(k-i))} \right) (k-i)^{\alpha/2}.$$

**Step 5 :** Regularity in time within intervals

For the regularity in time, we first consider  $t, s > 0$  with  $t+s \leq t_{i+1}, t \in ]t_i, t_{i+1}[$ . Then we use a variant of the composition formula for  $\lambda_t$  to get

$$\begin{aligned} h_t(x_1, \dots, x_i, x) &= \inf_{U \in L_a^2(\gamma_{[t, t+s]}, \mathbb{H})} \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [h_{t+s}(x_1, \dots, x_i, x + [(B - B_t) + U]_{t+s}) + \frac{1}{2}\|U\|_{\mathbb{H}}^2] \right) \right] \\ &\leq \left[ \mathbf{E}_{\gamma_{[t,1]}} ([h_{t+s}(x_1, \dots, x_i, x + (B_{t+s} - B_t))]) \right] \\ &\leq h_{t+s}(x_1, \dots, x_i, x) + \\ &\left[ \mathbf{E}_{\gamma_{[t,1]}} \left( (C_{k,i} \|(x_1, \dots, x_i, x)\| + C_{k,i} \|(x_1, \dots, x_i, x + (B_{t+s} - B_t))\| + D_{k,i}\sqrt{d}) \|(0, \dots, 0, B_{t+s} - B_t)\| \right) \right] \\ &\leq \left[ h_{t+s}(x_1, \dots, x_i, x) + (2C_{k,i} \|(x_1, \dots, x_i, x)\| + (C_{k,i} + D_{k,i})\sqrt{d})\sqrt{ds} \right]. \end{aligned}$$

Conversely, one obtains in considering  $U$  and noting that by Cauchy-Schwarz for  $U \in L_a^2(\gamma_{[t, t+s]}, \mathbb{H})$  we have  $\|U_{t+s}\| \leq \sqrt{s}\|U\|_{\mathbb{H}}$  and also using (3.3), (3.7) for  $U$  achieving enough the infimum of the left hand side:

$$\begin{aligned} \mathbf{E}_{\gamma_{[t,1]}} \left( h_{t+s}(x_1, \dots, x_i, x + [(B - B_t) + U]_s) + \frac{1}{2}\|U\|_{\mathbb{H}}^2 \right) &\geq [h_{t+s}(x_1, \dots, x_i, x) - \\ \mathbf{E}_{\gamma_{[t,1]}} \left( (C_{k,i} \|(B_{t+s} - B_t) + U_{t+s}\| + 2C_{k,i} \|(x_1, \dots, x_i, x)\| + D_{k,i}\sqrt{d}) \|B_{t+s} - B_t + U_{t+s}\| \right)] \\ &\geq \left[ h_{t+s}(x_1, \dots, x_i, x) - C_{k,i} \mathbf{E}_{\gamma_{[t,1]}} (\|B_{t+s} - B_t\| + \sqrt{s}\|U\|_{\mathbb{H}})^2 \right] \\ &\quad - (2C_{k,i} \|(x_1, \dots, x_i, x)\| + D_{k,i}\sqrt{d}) \sqrt{\mathbf{E}_{\gamma_{[t,1]}} (\|B_{t+s} - B_t\| + \sqrt{s}\|U\|_{\mathbb{H}})^2} \\ &\geq h_{t+s}(x_1, \dots, x_i, x) - 2C_{k,i}s(d + 2\epsilon + A(x_1, \dots, x_i, x)) \\ &\quad - 2(2C_{k,i} \|(x_1, \dots, x_i, x)\| + D_{k,i}\sqrt{d}) \sqrt{s(2\epsilon + A(x_1, \dots, x_i, x))}. \end{aligned}$$

Taking an infimum concludes to (3.5) in the present case  $t, s > 0, t > t_i$  with  $t+s \leq t_{i+1}$ .

**Step 6 :** Regularity in time of increments and derivatives within intervals

We are now ready to estimate time variation of increments for  $g \in \mathcal{E}_2(\mathbb{R}^{dk})$ . From  $h_{t+s} \in \mathcal{E}_2(\mathbb{R}^{dk})$ , we know that  $h_{t+s}$  is convex and locally Lipschitz and for  $x, z = x + y, t = x - y \in B(0, R)$ :

$$h_{t+s}(x+y) + h_{t+s}(x-y) - 2h_{t+s}(x) \leq (C_{\alpha,k,i} + D_{\alpha,k,i} \frac{3R}{\sqrt{d}}) \|y\|^2.$$

Thus if  $K = C_{\alpha,k,i} + D_{\alpha,k,i} \frac{3R}{\sqrt{d}}$ , we have by the parallelogram identity (since  $y = \frac{z-t}{2}$ ):

$$\begin{aligned} h_{t+s}(z) - \frac{K}{2} \|z\|^2 + h_{t+s}(t) - \frac{K}{2} \|t\|^2 - 2h_{t+s}\left(\frac{z+t}{2}\right) + K \left\| \frac{z+t}{2} \right\|^2 \\ \leq K \|y\|^2 - \frac{K}{2} \|z\|^2 - \frac{K}{2} \|t\|^2 + K \left\| \frac{z+t}{2} \right\|_2^2 = 0 \end{aligned}$$

so that  $\frac{K}{2} \|\cdot\|^2 - h_{t+s}$  is convex (since midpoint convex and continuous) on  $B(0, R)$ , thus one deduces from proposition 2.12 that  $h_{t+s}$  is Gateaux-differentiable with  $K$ -Lipschitz derivative

so that we can compute in applying the fundamental theorem of calculus along lines for  $Y, X + h, X, Y + h \in B(0, R)$ :

$$\begin{aligned}
 & |h_{t+s}(Y) + h_{t+s}(X + h) - h_{t+s}(X) - h_{t+s}(Y + h)| \\
 &= \left| \int_0^1 d\lambda dh_{t+s}(X + \lambda(Y - X)).(Y - X) - dh_{t+s}(X + \lambda(Y - X) + h).(Y - X) \right| \\
 (3.8) \quad &\leq K \|Y - X\| \|h\| \\
 &\leq (C_{\alpha,k,i} + 3D_{\alpha,k,i} \frac{\|Y\| + \|X\| + \|h\|}{\sqrt{d}}) \|Y - X\| \|h\|
 \end{aligned}$$

(with the last inequality obtained in minimizing  $R$ ).

Take a  $U$  giving a value  $\epsilon$ -close to the infimum in the formula for  $h_t(x_1, \dots, x_i, y)$ . We first take in the infimum definition for  $h_t(x_1, \dots, x_i, x)$  the value given at this  $U$ , apply the estimate just obtained and finally Cauchy-Schwarz inequality and the bound (3.7). We obtain for  $X = (x_1, \dots, x_i, x), Y = (x_1, \dots, x_i, y)$  :

$$\begin{aligned}
 h_t(X) + h_{t+s}(Y) &\leq h_{t+s}(Y) + \\
 &\left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [h_{t+s}(x_1, \dots, x_i, x + [(B - B_t) + U]_{t+s}) + \frac{1}{2} \|U\|_{\mathbb{H}}^2] \right) \right] \\
 &\leq h_{t+s}(X) + \mathbf{E}_{\gamma_{[t,1]}} \left( [h_{t+s}(x_1, \dots, x_i, y + [(B - B_t) + U]_{t+s}) + \frac{1}{2} \|U\|_{\mathbb{H}}^2] \right) \\
 &+ \mathbf{E}_{\gamma_{[t,1]}} \left( (C_{\alpha,k,i} + 3D_{\alpha,k,i} \frac{\|Y\| + \|X\| + \|[B - B_t] + U\|_{t+s}}{\sqrt{d}}) \|Y - X\| \|[B - B_t] + U\|_{t+s} \right) \\
 &\leq h_{t+s}(x_1, \dots, x_i, x) + h_t(x_1, \dots, x_i, y) + \epsilon \\
 &+ \|y - x\| (C_{\alpha,k,i} + 3D_{\alpha,k,i} \frac{\|(x_1, \dots, x_i, y)\| + \|(x_1, \dots, x_i, x)\|}{\sqrt{d}}) \mathbf{E}_{\gamma_{[t,1]}} ([\|B_{t+s} - B_t\| + \sqrt{s}\|U\|_{\mathbb{H}}]^2)^{1/2} \\
 &+ \|y - x\| \frac{3D_{\alpha,k,i}}{\sqrt{d}} \mathbf{E}_{\gamma_{[t,1]}} ([\|B_{t+s} - B_t\| + \sqrt{s}\|U\|_{\mathbb{H}}]^2) \\
 &\leq h_{t+s}(x_1, \dots, x_i, x) + h_t(x_1, \dots, x_i, y) + \epsilon \\
 &+ \|y - x\| (C_{\alpha,k,i} + 3D_{\alpha,k,i} \frac{\|(x_1, \dots, x_i, y)\| + \|(x_1, \dots, x_i, x)\|}{\sqrt{d}}) [2sd + 4\epsilon s + 2sA(x_1, \dots, x_i, y)]^{1/2} \\
 &+ \|y - x\| \frac{3D_{\alpha,k,i}}{\sqrt{d}} [2sd + 4\epsilon s + 2sA(x_1, \dots, x_i, y)].
 \end{aligned}$$

letting  $\epsilon \rightarrow 0$  and exchanging  $x, y$  one obtains the bound on increments :

$$\begin{aligned}
 (3.9) \quad & |h_t(x_1, \dots, x_i, x) - h_t(x_1, \dots, x_i, y) - h_{t+s}(x_1, \dots, x_i, x) + h_{t+s}(x_1, \dots, x_i, y)| \leq \|y - x\| \sqrt{s} \\
 &\times (C_{\alpha,k,i} + 3D_{\alpha,k,i} \frac{\|(x_1, \dots, x_i, y)\| + \|(x_1, \dots, x_i, x)\| + [2d + 2\max(A(x_1, \dots, x_i, y), A(x_1, \dots, x_i, x))]^{1/2}}{\sqrt{d}}) \\
 &\times [2d + 2\max(A(x_1, \dots, x_i, y), A(x_1, \dots, x_i, x))]^{1/2}
 \end{aligned}$$

and since  $h_t$  admits partial derivatives, in making in making  $x \rightarrow y$ , one obtains for  $\|H\| = 1, H \in \mathbb{R}^d$ , a bound on the directional derivative  $D_H^y h_t(x_1, \dots, x_i, y) = \sum_{i=1}^d \frac{\partial}{\partial y^{(i)}} h_t(x_1, \dots, x_i, y) H_j$ :

$$\begin{aligned}
 & |D_H^y h_t(x_1, \dots, x_i, y) - D_H^y h_{t+s}(x_1, \dots, x_i, y)| \\
 &\leq s^{1/2} (C_{\alpha,k,i} + 3D_{\alpha,k,i} \frac{2\|(x_1, \dots, x_i, y)\| + [2d + 2A(x_1, \dots, x_i, y)]^{1/2}}{\sqrt{d}}) [2d + 2A(x_1, \dots, x_i, y)]^{1/2}.
 \end{aligned}$$

**Step 7 :** Regularity in time around a  $t_i$



For the second and last case,  $t = t_i, s > 0, t + s < t_{i+1}$

$$\begin{aligned}
h_t(x_1, \dots, x_i) &= \inf_{U \in L_a^2(\gamma_{[t,s]}, \mathbb{H})} \left[ \mathbf{E}_{\gamma_{[t,1]}} \left( [h_{t+s}(x_1, \dots, x_i, x_i + [(B - B_t) + U]_{t+s}) + \frac{1}{2} \|U\|_{\mathbb{H}}^2] \right) \right] \\
&\leq \left[ \mathbf{E}_{\gamma_{[t,1]}} ([h_{t+s}(x_1, \dots, x_i, x_{i+1} + (x_i - x_{i+1}) + (B_{t+s} - B_t))]) \right] \\
&\leq h_{t+s}(x_1, \dots, x_i, x_{i+1}) \\
&\quad + (2C_{k,i} \|(x_1, \dots, x_i, x_{i+1})\| + (C_{k,i} + D_{k,i})(\sqrt{d} + \|x_i - x_{i+1}\|))(\sqrt{ds} + \|x_i - x_{i+1}\|).
\end{aligned}$$

and similarly:

$$\begin{aligned}
h_t(x_1, \dots, x_i) &\geq h_{t+s}(x_1, \dots, x_i, x) - 3C_{k,i} (\|x_i - x_{i+1}\| + sd + sA(x_1, \dots, x_i, x_i)) \\
&\quad - 3(2C_{k,i} \|(x_1, \dots, x_i, x_{i+1})\| + D_{k,i}\sqrt{d})\sqrt{sA(x_1, \dots, x_i, x_i) + \|x_i - x_{i+1}\|}.
\end{aligned}$$

□

#### 4. STOCHASTIC DIFFERENTIAL EQUATIONS WITH MONOTONE DRIFT AND THEIR FREE VARIANT

**4.1. Classical Case.** We quote here the main result of the chapter 3 in [PR] (coming from [Kr]) and apply it to the setting we need.

We start by quoting their Theorem 3.1.1. We consider  $W_t$  a Wiener process in  $\mathbb{R}^d$  in a normal filtration  $\mathcal{F}_t$  (i.e. for instance the completed filtration generated by this brownian motion on Wiener space  $\Omega$ , to insure  $\mathcal{F}_0$  contains every null-sets and the filtration is right continuous, cf. [PR, Prop 2.1.13]). We fix

$$\sigma : [0, 1] \times \mathbb{R}^d \times \Omega \rightarrow M_d(\mathbb{R}), \quad b : [0, 1] \times \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}^d$$

continuous in  $x \in \mathbb{R}^d$  for  $t \in [0, 1], w \in \Omega$  fixed and progressively measurable in the sense that their restriction to  $[0, t] \times \Omega$  is  $B([0, t]) \otimes \mathcal{F}_t$  measurable. We give the target space their usual euclidean norms  $\|\cdot\|$ .

**Theorem 4.1** ([PR] Theorem 3.1.1 ). *Consider  $b, \sigma$  as above, and assume moreover that on  $\Omega$  for all,  $R \in [0, \infty[$  we have the integrability condition:*

$$(4.1) \quad \int_0^1 dt \sup_{|x| \leq R} \|\sigma(t, x)\|^2 + \|b(t, x)\| < \infty$$

and for also  $t \in [0, 1], x, y \in \mathbb{R}^d$ , with  $\|x\|, \|y\| \leq R$  the local weak monotonicity:

$$(4.2) \quad 2\langle x - y, b(t, x) - b(t, y) \rangle + \|\sigma(t, x) - \sigma(t, y)\|^2 \leq K_t(R)\|x - y\|^2$$

and the weak coercivity:

$$(4.3) \quad 2\langle x, b(t, x) \rangle + \|\sigma(t, x)\|^2 \leq K_t(1)(d + \|x\|^2)$$

where for each  $R > 0$ ,  $K_t(R)$  is an  $\mathbb{R}_+$ -valued,  $\mathcal{F}_t$  adapted process satisfying  $\alpha_s(R) = \int_0^s K_t(R)dt < \infty$  on  $\Omega$ . Then, Then for any  $\mathcal{F}_0$ -measurable map  $X_0 : \Omega \rightarrow \mathbb{R}^d$  there exists a (up to  $P$ -indistinguishability) unique solution to the stochastic differential equation

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dB(t).$$

Here solution means that  $(X(t))_{t \geq 0}$  is a  $P$ -a.s. continuous  $\mathbb{R}^d$ -valued  $\mathcal{F}_t$ -adapted process such that  $P$ -a.s. for all  $t \in [0, 1]$

$$(4.4) \quad X(t) = X_0 + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dB(s).$$

Furthermore, for all  $t \in [0, 1]$

$$E(\|X(t)\|^2 e^{-\alpha t(1)}) \leq E(\|X_0\|^2) + d.$$

We now apply this result in the simple case we will use in taking care of dimensional dependence of constants as before.

**Corollary 4.2.** *We consider the previous setting with  $\Omega = C^0([0, 1], \mathbb{R}^d)$  the pathspace with Wiener measure  $\gamma$  with its canonical normal filtration. Let  $g \in \mathcal{E}_2(\mathbb{R}^{dk})$  and  $t_0 = 0 \leq t_1 < \dots < t_k \in [0, 1]$  and  $h_t$  defined in lemma 3.5. We define  $b(t, x, \omega)$  for  $\omega \in \Omega$ , for  $t_i < t \leq t_{i+1}$  by :*

$$b_j(t, x, \omega) = b_j^g(t, x, \omega) := -\frac{\partial}{\partial x^{(j)}} h_t(X_{t_1}(\omega), \dots, X_{t_i}(\omega), x),$$

and  $b(t, x, \omega) = 0$  si  $t > t_k$ . Fix also  $\sigma(t, x, \omega) = I_d$ . Then, they satisfy (4.1), (4.2) and (4.3) for  $K_t(R) = 0$  for  $R \neq 1$ ,  $K_t(1) = K > 0$  a fixed constant so that the conclusion of theorem (4.1) follows on each interval, with a solution  $X_s$  which is then used to define  $b$  on the next interval. Moreover, for the solution  $X_s$ , we have the regularity bounds for some  $C_4$  for all  $t, s \in ]t_i, t_{i+1}], t > s$ :

$$E(\|b(t, X_t) - b(s, X_s)\|) \leq C_4 \sqrt{d(t-s)} \left( 1 + \frac{E(\|X_0\|^2) + \sup(-g)}{d} \right),$$

$$E(\|b(t, X_t) - b(s, X_s)\|^2) \leq C_4 d \sqrt[4]{(t-s)} \sqrt{\left( 1 + \frac{E(\|X_0\|^2) + \sup(-g)}{d} \right)} \left( 1 + \frac{E(\|X_0\|^4)}{d^2} \right)^{3/8},$$

and the estimate for  $t \in [0, 1]$ :

$$E(\|X(t)\|^4) \leq [E(\|X(0)\|^4) + (3K + 1)d^2] e^{(3K+1)t}.$$

The last Hölder-continuity estimate will be crucial to recover some continuity in ultraproducts.

*Proof.* By induction, one can assume the previously built solution  $X_t$  to be measurable. From (3.4) and proposition 2.12,  $\frac{\partial}{\partial x^{(j)}} h_t(X_{t_1}(\omega), \dots, X_{t_i}(\omega), x)$  is Lipschitz-continuous in  $x$ , thus continuous. From proposition 3.5, it is also continuous in  $t$  on  $]t_i, t_{i+1}]$  for each  $\omega, x$  fixed and for each  $x, t \in [0, 1]$  the formula is clearly  $\sigma(X_s, s \leq t)$ -measurable thus, by e.g. [M, lemma 9.2], on  $]t_i, t_{i+1}] \times \Omega$ ,  $b_j(\cdot, x, \cdot)$  is  $B([t_i, t]) \otimes \mathcal{F}_t$ -measurable, and thus  $b$  is progressively measurable. We have from (3.3):

$$\sup_{\|x\|_2 \leq R} \|\sigma(t, x)\|^2 + \|b(t, x)\| \leq 1 + (2CR + D\sqrt{d}),$$

so that (4.1) is verified. The monotony of  $b(t, \cdot, \omega)$  follows from the convexity of  $h$  so that (4.2) holds with  $K_t(R) = 0$ . (4.3) then follows from (3.3) as before since  $\langle x, b(t, x, \omega) \rangle = dh_t(\omega_{t_1}, \dots, \omega_{t_i}, x) \cdot (x_1, \dots, x_n)$  with  $K_t(1) = \max(5C, D^2) + 1$ .

For the continuity property, we decompose the inequality into two terms. First a bound on  $E(\|b(t, X_t) - b(s, X_t)\|)$  is obtained by using the last inequality in proposition 3.5 and the estimates on  $E(\|X_t\|^2)$ .

$$E(\|b(t, X_t) - b(s, X_t)\|) \leq \sqrt{|t-s|d} \left( c_1 + d_1 \frac{e^K E(\|X_0\|^2 + d) + \sup(-g)}{d} \right).$$

A similar bound holds from the application of proposition 2.12 as in step 6 of the proof of proposition 3.5:

$$\begin{aligned} E(\|b(s, X_t) - b(s, X_s)\|) &\leq E \left[ (C + D \frac{\|X_t\| + \|X_s\|}{\sqrt{d}}) \|X_t - X_s\| \right] \\ &\leq (C + D \frac{\sqrt{E(\|X_t\|^2)} + \sqrt{E(\|X_s\|^2)}}{\sqrt{d}}) \sqrt{E(\|X_t - X_s\|^2)}, \end{aligned}$$

and since, we have:  $E(\|b(u, X_u)\|^2) \leq E(2C\|X_u\| + D\sqrt{d})^2 \leq 8C^2e^K(E(\|X_0\|^2) + d) + 2D^2d$ , and say for  $t > s$

$$\begin{aligned} E(\|X_t - X_s\|^2) &\leq 2d(t-s) + 2\sqrt{d(t-s)}\sqrt{E(\int_s^t \|b(u, X_u)\|^2 du)} \\ &\leq 2d(t-s) + 2\sqrt{d(t-s)}\sqrt{8C^2e^K(E(\|X_0\|^2) + d) + 2D^2d} = C_3d(t-s). \end{aligned}$$

Thus, we obtain the expected bound:

$$E(\|b(s, X_t) - b(s, X_s)\|) \leq (C + 2De^{K/2} \frac{\sqrt{E(\|X_0\|^2) + d}}{\sqrt{d}}) \sqrt{C_3d(t-s)}.$$

The fourth order estimate is not contained in the previous stated result (but it is well-known, see e.g. [BL, Th 3.5 p 59]) Ito's formula and our weak coercivity assumptions gives :

$$\begin{aligned} \|X(t)\|^2 &= \|X(0)\|^2 + \int_0^t (2\langle X(s), b(s, X(s)) \rangle + 1)ds + \int_0^t \langle X(s), dB_s \rangle \\ &\leq \|X(0)\|^2 + K \int_0^t (1 + \|X(s)\|^2)ds + \int_0^t \langle X(s), dB_s \rangle \end{aligned}$$

In taking the square and applying Ito's formula again and taking mathematical expectation, one gets:

$$\begin{aligned} E(\|X(t)\|^4) &\leq E\left(\|X(0)\|^4 + 2K \int_0^t (d + \|X(s)\|^2)\|X(s)\|^2 ds + \int_0^t \|X(s)\|^2 ds\right) \\ &\leq E\left(\|X(0)\|^4 + (3K + 1) \int_0^t (d^2 + \|X(s)\|^4) ds\right) \end{aligned}$$

Gronwall's lemma concludes to the fourth order bound and then Cauchy-Schwartz inequality gives the  $L^2$  bound on increments:

$$\begin{aligned} E(\|b(t, X_t) - b(s, X_s)\|^2) &\leq E(\|b(t, X_t) - b(s, X_s)\|^3)^{1/2} E(\|b(t, X_t) - b(s, X_s)\|)^{1/2} \\ &\leq 4E(\|b(t, X_t) - b(s, X_s)\|)^{1/2} \sup_{t \in [0,1]} E(\|b(t, X_t)\|^4)^{3/8}. \end{aligned}$$

□

**4.2. Free case.** We now obtain a result similar to the previous corollary in the free case in order to describe the limit of the application to matrices of this corollary. Instead of using Euler approximation as in [PR], we will use a Yosida approximation as in [LS, Theorem 4.3]. The necessary preliminaries were recalled in subsection 2.6. The following result is of independent interest for the study of free SDEs. That's why we assume a slightly more general setting that what we need later.

**Theorem 4.3.** *Let  $M_t \subset (M, \tau)$  a filtration of finite von Neumann algebras containing an adapted free brownian motion  $S_t = (S_t^1, \dots, S_t^m)$ . Let  $T > 0$  and  $t_0 = 0 < t_1 < t_2 \dots < t_k \leq t_{k+1} = T$   $|t_{i+1} - t_i| \leq 1$ , and for  $t \in [t_i, t_{i+1}]$ , let  $h_t : L^2(M_t, \tau)^{m(i+1)} \rightarrow \mathbb{R}$  a convex function bounded below by  $c \in \mathbb{R}$  (uniformly in  $t$ ), subquadratic with bound:*

$$h_t(x) \leq |c|(1 + \|x\|_2^2),$$

and satisfying for some  $C, D > 0$ , for  $X, Y \in L^2(M_t, \tau)^{m(i+1)}$ :

$$|h_t(X) - h_t(Y)| \leq \|X - Y\| (C\|X\|_2 + C\|Y\|_2 + D).$$

Assume  $h_t$  is Gâteaux differentiable such that  $t \mapsto \nabla_{i+1} h_t(X)$  is continuous with value  $L^2(M, \tau)$  on any  $[t_i, t_{i+1}]$ , that for  $X \in L^2(M_u)^{m(i+1)}$ ,  $u < t$  we have  $\nabla_{i+1} h_t(X) \in L^2(M_u)$  and even satisfying for some  $\alpha, \beta \in [0, 1]$  and any  $t < s \in [t_i, t_{i+1}]$ ,  $X, Y \in L^2(M_t, \tau)^{m(i+1)}$ :

$$(4.5) \quad \|\nabla_{i+1} h_t(X) - \nabla_{i+1} h_s(X)\|_2 \leq |t - s|^\alpha (C + D\|X\|_2).$$

$$(4.6) \quad \|\nabla_{i+1} h_t(X) - \nabla_{i+1} h_t(Y)\|_2 \leq \|X - Y\|_2^\beta (C + D\|X\|_2 + D\|Y\|_2).$$

Assume finally given  $X_0 \in L^2(M_0, \tau)$ . Then there is  $X_t \in L^2(M_t, \tau)$  continuous in  $t$  satisfying :

$$X_t = X_0 + S_t - \int_0^t u_s(X) ds, \quad u_s(X) = \sum_{i=0}^k \nabla_{i+1} h_s(X_{t_1}, \dots, X_{t_i}, X_s) 1_{]t_i, t_{i+1}]}(s).$$

Moreover, any continuous solution  $Y_t \in L^2(M_t, \tau)$  of this equation equals  $X_t$  for every  $t \in [0, T]$  if  $Y_0 = X_0$ .

*Proof.* **Step 1 :** Uniqueness

Let us start by checking uniqueness. Note that by the assumption  $\|u_s(X)\|_2 \leq (2C\|(X_{t_1}, \dots, X_{t_i}, X_s)\|_2 + D)$  which is bounded on  $[0, T]$  and

$$X_t - Y_t = X_0 - Y_0 - \int_0^t (u_s(X) - u_s(Y)) ds = - \int_0^t (u_s(X) - u_s(Y)) ds$$

so that  $X_t - Y_t$  is absolutely continuous in  $t$  in  $L^2(M)$  and we have:

$$\|X_t - Y_t\|_2^2 = -2 \int_0^t \langle (u_s(X) - u_s(Y)), X_s - Y_s \rangle \leq 0$$

since  $\langle (u_s(X) - u_s(Y)), X_s - Y_s \rangle \geq 0$  since  $u_s$  is monotone from the convexity of  $h_s$  and first since for  $s \in [0, t_1]$   $u_s(X) \in L^2(M_s)$  and then since by induction on the time intervals for  $s \in ]t_i, t_{i+1}]$  one can use  $X_{t_i} = Y_{t_i}$ .

**Step 2 :** Definition of the Yosida approximation

We now turn to the proof of the existence result. We follow a Yosida approximation scheme (see e.g. [LS, Theorem 4.3] in the classical case). By induction on  $i$ , we aim at finding a solution on  $]t_i, t_{i+1}]$  and thus we consider given  $X_{t_1}, \dots, X_{t_i}$  and  $H_t = h_t(X_{t_1}, \dots, X_{t_i}, \cdot) : L^2(M_t) \rightarrow \mathbb{R}$ . This is a convex continuous (even locally Hölder continuous) function and we can consider  $H_{t,\lambda}$  its Hopf-Lax-Yosida approximation from Proposition 2.14. We now fix  $0 < \lambda \leq 1$ . From a Picard iteration argument (see e.g. [Ga, Lemma 3.2], or the proof of Theorem 6.2 Step 2(ix) in a more complicated context), since  $\nabla H_{t,\lambda}$  is globally  $1/\lambda$ -Lipschitz, one gets a solution on  $]t_i, t_{i+1}]$  with  $X_{t,\lambda} \in L^2(M_t)$ :

$$X_{t,\lambda} = X_{t_i} + S_t - S_{t_i} - \int_{t_i}^t ds \nabla H_{s,\lambda}(X_{s,\lambda}).$$

Note that this argument uses the regularity in time obtained from lemma 2.15 based on the assumptions (4.6) and (4.5) so that the Picard iterates are well-defined and adapted. Note that using the assumption that for  $X \in L^2(M_u)^{m(i+1)}$ ,  $u < t$  we have  $\nabla_{i+1} h_t(X) \in L^2(M_u)$ , the quoted lemma can be used with  $H = L^2(M_t)$  and the appropriate restriction to this space of  $h_s, s > t$ . Even in the case where  $\beta = 1$ , the Yosida approximation is useful to get a global lipschitzness.

**Step 3 :** Second moment estimate

From Ito's formula [BS] and taking traces, one gets the a priori estimate:

$$\begin{aligned}
\|X_{t,\lambda}\|_2^2 &= \|X_{t_i}\|_2^2 - 2 \int_{t_i}^t ds \langle \nabla H_{s,\lambda}(X_{s,\lambda}), X_{s,\lambda} \rangle + (t - t_i) \\
&\leq \|X_{t_i}\|_2^2 + 2 \int_{t_i}^t ds (2C\|(X_{t_1}, \dots, X_{t_i}, X_{s,\lambda})\|_2 + 2C\|J_{s,\lambda}(0)\|_2 + D)\|X_{s,\lambda}\|_2 + (t - t_i) \\
&\leq \|X_{t_i}\|_2^2 + 2 \int_{t_i}^t ds (2C + D + 4C(|c| + H_s(0)) + 2C\|(X_{t_1}, \dots, X_{t_i})\|_2)\|X_{s,\lambda}\|_2^2 \\
&\quad + (D/4 + C(2|c| + |c|\|(X_{t_1}, \dots, X_{t_i})\|_2^2) + C\|(X_{t_1}, \dots, X_{t_i})\|_2 + 1)(t_{i+1} - t_i) \\
&\leq (\|X_{t_i}\|_2^2 + (1 + C(X_{t_1}, \dots, X_{t_i})/4)(t_{i+1} - t_i)) e^{C(X_{t_1}, \dots, X_{t_i})(t-t_i)} =: E(t),
\end{aligned}$$

where the inequality comes from the lipschitzness of  $h_t$  and  $\nabla H_{s,\lambda} = \nabla H_s J_{s,\lambda}$  for the canonical contraction  $J_{s,\lambda} = (I + \lambda \nabla H_s)^{-1}$  satisfying the bound (2.13). The last inequality comes from Gronwall's lemma and we defined  $C(X_{t_1}, \dots, X_{t_i}) = 2(2C + D + 4C(2|c| + |c|\|(X_{t_1}, \dots, X_{t_i})\|_2^2) + 2C\|(X_{t_1}, \dots, X_{t_i})\|_2)$ .

**Step 4 :** Convergence in  $\lambda$  in  $L^2$ .

Then, one deduces similarly for  $t \in [t_i, t_{i+1}]$ :

$$\|X_{t,\lambda} - X_{t,\mu}\|_2^2 = -2 \int_{t_i}^t ds \langle \nabla H_{s,\lambda}(X_{s,\lambda}) - \nabla H_{s,\mu}(X_{s,\mu}), X_{s,\lambda} - X_{s,\mu} \rangle$$

Now, note that, from proposition 2.14, if we call  $J_{s,\lambda} = (I + \lambda \nabla H_s)^{-1}$  the contraction, we have  $\nabla H_{s,\lambda} = \nabla H_s \circ J_{s,\lambda}$  and  $\lambda \nabla H_{s,\lambda}(x) + J_{s,\lambda}(x) = x$ , thus we can decompose

$$\begin{aligned}
&\langle \nabla H_{s,\lambda}(X_{s,\lambda}) - \nabla H_{s,\mu}(X_{s,\mu}), X_{s,\lambda} - X_{s,\mu} \rangle \\
&= \langle \nabla H_s(J_{s,\lambda}(X_{s,\lambda})) - \nabla H_s(J_{s,\mu}(X_{s,\mu})), J_{s,\lambda}(X_{s,\lambda}) - J_{s,\mu}(X_{s,\mu}) \rangle \\
&\quad + \langle \nabla H_{s,\lambda}(X_{s,\lambda}) - \nabla H_{s,\mu}(X_{s,\mu}), \lambda \nabla H_{s,\lambda}(X_{s,\lambda}) - \mu \nabla H_{s,\mu}(X_{s,\mu}) \rangle \\
&\geq \langle \nabla H_{s,\lambda}(X_{s,\lambda}) - \nabla H_{s,\mu}(X_{s,\mu}), \lambda \nabla H_{s,\lambda}(X_{s,\lambda}) - \mu \nabla H_{s,\mu}(X_{s,\mu}) \rangle
\end{aligned}$$

where we use the convexity of  $H$  in the inequality. Thus one gets in using more the relations above (and the inequality  $|ab| \leq a^2/4 + b^2$ ) :

$$\begin{aligned}
&-\langle \nabla H_{s,\lambda}(X_{s,\lambda}) - \nabla H_{s,\mu}(X_{s,\mu}), X_{s,\lambda} - X_{s,\mu} \rangle \\
&\leq -\lambda \|\nabla H_{s,\lambda}(X_{s,\lambda})\|^2 - \mu \|\nabla H_{s,\mu}(X_{s,\mu})\|^2 + (\lambda + \mu) \|\nabla H_{s,\lambda}(X_{s,\lambda})\| \|\nabla H_{s,\mu}(X_{s,\mu})\| \\
&\leq \frac{\lambda}{4} \|\nabla H_{s,\mu}(X_{s,\mu})\|^2 + \frac{\mu}{4} \|\nabla H_{s,\lambda}(X_{s,\lambda})\|^2.
\end{aligned}$$

Finally, from the Lipschitzness of  $H$ , one gets from the contractivity of  $J_{s,\mu}$ :

$$\begin{aligned}
\|\nabla H_{s,\mu}(X_{s,\mu})\|^2 &\leq (2C\|(X_{t_1}, \dots, X_{t_i}, J_{s,\mu}(X_{s,\mu}))\| + D)^2 \\
&\leq (2C\|(X_{t_1}, \dots, X_{t_i}, X_{s,\mu})\| + 2C\|J_{s,\mu}(0)\| + D)^2
\end{aligned}$$

Gathering all our estimates and using again (2.13), we have thus obtained the inequality (for  $\mu, \lambda \leq 1$ ):

$$\begin{aligned}
\|X_{t,\lambda} - X_{t,\mu}\|_2^2 &\leq \frac{\lambda + \mu}{2} (t - t_i) \\
&\quad \times (8C^2 \sum_{j=1}^i \|X_{t_j}\|_2^2 + 8C^2 E(t_{i+1}) + 2(D + 4C(2|c| + |c|\|(X_{t_1}, \dots, X_{t_i})\|_2^2))^2).
\end{aligned}$$

Thus  $X_{t,\lambda}$  is Cauchy in  $\lambda$  and converges in  $C^0([t_i, t_{i+1}], L^2(M))$  to  $X_t$  such that  $X_t \in L^2(M_t)$ .

**Step 5 :** Checking the limiting equation.

Moreover, using (4.6), triangular inequality and the definition of  $J_{s,\mu}$ , one gets:

$$\begin{aligned} \|\nabla H_{s,\mu}(X_{s,\mu}) - \nabla H_s(X_s)\|_2 &= \|\nabla H_s J_{s,\mu}(X_{s,\mu}) - \nabla H_s(X_s)\|_2 \\ &\leq \|J_{s,\mu}(X_{s,\mu}) - X_s\|_2^\beta (C + D\|J_{s,\mu}(X_{s,\mu})\|_2 + D\|X_s\|_2) \\ &\leq \left( \mu^\beta \|\nabla H_{s,\mu}(X_{s,\mu})\|_2^\beta + \|X_{s,\mu} - X_s\|_2^\beta \right) (C + D\|X_{s,\mu}\|_2 + D\|J_{s,\mu}(0)\|_2 + D\|X_s\|_2) \rightarrow_{\mu \rightarrow 0} 0 \end{aligned}$$

uniformly on  $[t_i, t_{i+1}]$  from the bounds in the previous steps. We can now take the limit in the SDE satisfied by  $X_{s,\mu}$  to get the expected SDE for  $X_s$ .  $\square$

We now apply this result in the spirit of [GS09] in order to get limit states of convex potential matrix models with limited regularity.

**Theorem 4.4.** *Let  $g \in \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_1^m * \mathcal{F}_\mu^\nu), d_{2,0})$  and consider, for  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))$ , the law absolutely continuous with respect to the law  $P_{G^N}$  of GUE  $G^N$ :*

$$d\mu_{g,N}(X) = \frac{1}{Z_{g,N,\Upsilon_N}} e^{-N^2 g(\tau_X, \Upsilon_N)} dP_{G^N}(X).$$

Assume finally that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\Upsilon \in (\mathcal{T}(\mathcal{F}_\mu^\nu), d)$ . Then  $E_{\mu_{g,N}} \circ \tau_{\Upsilon_N}$  converges in  $(\mathcal{T}_{2,0}(\mathcal{F}_1^m * \mathcal{F}_\mu^\nu), d_{2,0})$  to a tracial state  $\tau_g$  which is law of self-adjoint and unitary variables  $X(g), u$  (of norm bounded by some  $R$ ) and the unique solution, such that the law of  $u$  is  $\mu_\Upsilon$ , to the equation  $(SD_g)$ , for  $G(X) = g(\tau_{X,u})$ :

$$\forall P \in \mathbb{C}\langle X_1, \dots, X_n, u_1^1, \dots, u_\mu^\nu \rangle, (\tau_g \otimes \tau_g)(\partial_{X_i}(P)) = \tau_g(X_i P) + d_{X_i} G(X(g)).P(X).$$

Moreover, there is a solution on  $\mathbb{R}_+$  given by the theorem 4.3 for  $h_t(x_1, \dots, x_{i+1}) = g(\tau_{x_{i+1},u}) + g_{2,(1)}(\tau)$  for all  $t$  (with  $(\mathbf{1})$  the list of times with only one time equal to 1), it satisfies for the solutions  $X_t(X), X_t(Y)$  with initial condition  $X, Y$ :

$$(4.7) \quad \|X_t(X) - X_t(Y)\|_2^2 \leq e^{-t} \|X_0(X) - X_0(Y)\|_2^2$$

and  $\tau_g$  is the unique stationary state for this free SDE.

*Proof. Step 1 :* Defining limit variables in a von Neumann algebra ultraproduct.

Let law  $\mu_{g,N}$  be the marginal at time 1 of a law considered in Proposition 2.7. Consider a non-principal ultrafilter  $\omega$  on  $\mathbb{N}$  and the ultraproducts  $\mathcal{L}^\omega = L^2(M_N(L^\infty(\mu_{g,N}))^\omega)$ ,  $\mathcal{M}^\omega = M_N(L^\infty(\mu_{g,N}))^\omega$  (tracial von Neumann algebra ultraproduct). Considering  $A_1^N, \dots, A_n^N$  the canonical hermitian variables in  $M_N(L^\infty(\mu_{g,N}))$ , we know from (2.8) that  $\|A_i^N 1_{\|A_i^N\| \leq C} - A_i^N\|_2 \rightarrow 0$  so that  $X_i^\omega = (A_i^N)^\omega = (A_i^N 1_{\|A_i^N\| \leq C})^\omega \in \mathcal{M}^\omega$ . We thus also fix  $B_i^N = A_i^N 1_{\|A_i^N\| \leq C}$ . We can also consider  $u_i^j = ((\Upsilon_N)_i^j)^\omega \in \mathcal{M}^\omega$ . Of course,  $u$  has law  $\mu_\Upsilon$ .

This gives a tracial state  $\tau_{X^\omega, u} \in \mathcal{S}_C^m \star \mathcal{T}_\mu^\nu$ . Let us check that any such state satisfies  $(SD_g)$ .

**Step 2 :** Showing  $(SD_g)$ .

First, note that for  $P_1, \dots, P_m \in \mathbb{C}\langle X_1, \dots, X_m, u_1^1, \dots, u_\mu^\nu \rangle$ ,

$$\begin{aligned} \lim_{N \rightarrow \omega} E_{\mu_{g,N}}(g(\tau_{B_1^N + P_1(B_1^N, \dots, B_m^N, \Upsilon_N), \dots, B_m^N + P_m(B_1^N, \dots, B_m^N, \Upsilon_N), \Upsilon_N})) \\ = G(X_1^\omega + P_1(X_1^\omega, \dots, X_m^\omega), \dots, X_m^\omega + P_m(X_1^\omega, \dots, X_m^\omega, u, u)). \end{aligned}$$

Indeed, from the lipschitzness of  $g$  for the metric  $d_2$  it is easy to see that

$$g(\tau_{X_1 + P_1(X_1, \dots, X_m), \dots, X_m + P_m(X_1, \dots, X_m), u}) =: G_P(\mu_{X,u})$$

is uniformly continuous on  $\mathcal{S}_C^m \star \mathcal{T}_\mu^\nu$  for the induced metric  $D_2(\mu, \nu) = d_{2,0}(\tau_{X(\mu), u(\mu)}, \tau_{X(\nu), u(\nu)})$  and from the second concentration (2.10) in Proposition 2.7,  $E_{\mu_{g,N}}(d_{2,0}(\tau_{B^N, \Upsilon_N}, E_{\mu_{g,N}}(\tau_{B^N, \Upsilon_N}))) \rightarrow_{N \rightarrow \infty} 0$ , so that  $\lim_{N \rightarrow \omega} E_{\mu_{g,N}}(d_{2,0}(\tau_{B^N, \Upsilon_N}, \tau_{X^\omega, u})) = 0$  and as a consequence for any  $\eta > 0$

$$\lim_{N \rightarrow \omega} P(d_{2,0}(\tau_{B^N, \Upsilon_N}, \tau_{X^\omega, u}) > \eta) = 0.$$



If we fix  $\epsilon > 0$  and find  $\eta > 0$  such that if  $d_{2,0}(\tau_{X(\mu),u(\mu)}, \tau_{X(\nu),u(\nu)}) \leq \eta$  then  $|G_P(\tau_{X(\mu),u(\mu)}) - G(\tau_{X(\nu),u(\nu)})| \leq \epsilon$  one deduces as claimed that:

$$\lim_{N \rightarrow \omega} \left| E_{\mu_{g,N}} \left( g(\tau_{B_1^N + P_1(B_1^N, \dots, B_m^N), \dots, B_m^N + P_m(B_1^N, \dots, B_m^N), \Upsilon_N}) \right) - G(X_1^\omega + P_1(X_1^\omega, \dots, X_m^\omega, u), \dots, X_n^\omega + P_n(X_1^\omega, \dots, X_m^\omega, u)) \right| \leq \epsilon + \lim_{N \rightarrow \omega} \sqrt{P(d_2(\tau_{B^N, \Upsilon_N}, \tau_{X^\omega, u}) > \eta)(c + E(G^2(B_1^N + P_1(B_1^N, \dots, B_m^N, \Upsilon_N), \dots, B_m^N + P_1(B_1^N, \dots, B_m^N, \Upsilon_N), \Upsilon_N))) = \epsilon$$

where the last equality comes from the previously found convergence in probability and the subquadratic behaviour of  $G$  in conjunction with  $\|B_i^N\| \leq C$  and  $c > 0$  is another constant.

As in [GM], we use an integration by parts formula on  $\mu_{g,N}$  which gives  $\forall P \in \mathbb{C}\langle X_1, \dots, X_m, u_1^1, \dots, u_\mu^\nu \rangle$ :

$$\begin{aligned} E_{\mu_{g,N}} \left( \frac{1}{N} \text{Tr}(A_i^N P(A_1^N, \dots, A_m^N, \Upsilon_N)) + \frac{1}{N} \text{Tr}(N \nabla_{A_i^N} G(A_1^N, \dots, A_m^N) P(A_1^N, \dots, A_m^N, \Upsilon_N)) \right) \\ = E_{\mu_{g,N}} \left( \left( \frac{1}{N} \text{Tr} \otimes \frac{1}{N} \text{Tr} \right) (\partial_{X_i} P)(A_1^N, \dots, A_m^N, \Upsilon_N) \right) \end{aligned}$$

and the second concentration result in Proposition 2.7 implies that the right hand side converges when  $N \rightarrow \omega$  to  $(\tau_{X^\omega, u} \otimes \tau_{X^\omega, u})(\partial_i(P))$ .

But from the lipschitzness condition (2.2) on  $G$ , one deduces

$$\text{Tr}(\nabla_{A_i^N} G(A_1^N, \dots, A_m^N) \nabla_{A_i^N} G(A_1^N, \dots, A_m^N)) \leq \frac{C}{N} (1 + \frac{1}{N} \text{Tr}(\sum_{i=1}^m (A_i^N)^2))$$

and thus if  $Z_i^N = N \nabla_{A_i^N} G(A_1^N, \dots, A_m^N)$  one gets  $E(\|Z_i^N\|_2^2) \leq C(1 + E(\frac{1}{N} \text{Tr}(\sum_{i=1}^m (A_i^N)^2)))$ , so that  $Z = (Z^N)^\omega \in \mathcal{L}^\omega$  and one obtains the relation in taking of limit to  $\omega$  of the integration by parts relation:

$$(4.8) \quad \langle X_i^\omega + Z_i^*, P(X_1^\omega, \dots, X_m^\omega, u) \rangle = (\tau_{X^\omega, u} \otimes \tau_{X^\omega, u})(\partial_{X_i}(P)).$$

But since from the definition of the gradient

$$\begin{aligned} |E_{\mu_{g,N}} (G(B_1^N + tP_1(B_1^N, \dots, B_m^N, \Upsilon_N), \dots, B_m^N + tP_m(B_1^N, \dots, B_m^N, \Upsilon_N) - G(B_1^N, \dots, B_m^N)) \\ - tE_{\mu_{g,N}} (\sum_{i=1}^m \frac{1}{N} \text{Tr}(Z_i^N P_1(B_i^N, \dots, B_m^N, \Upsilon_N)))| \leq t^2 \sum_{i=1}^m \frac{c}{N} \text{Tr}(P_i^2(B_1^N, \dots, B_m^N, \Upsilon_N)) \end{aligned}$$

Thus taking the limit  $N \rightarrow \Omega$  one deduces

$$\langle Z_i^*, P(X_1^\omega, \dots, X_m^\omega, u) \rangle = d_{X_i^\omega} G(X_1^\omega, \dots, X_m^\omega, u) \cdot P(X_1^\omega, \dots, X_m^\omega, u).$$

This shows (4.8) was the equation  $(SD_g)$  we were aiming at. Moreover, note that this implies  $\tau_{X^\omega, u}$  has finite Fisher information.

**Step 3 :** Properties and use of the SDE.

Since  $h_t$  does not depend on time, assumption (4.5) is obvious and all the remaining assumptions in Theorem 4.3 are contained in  $g \in \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_1^m * \mathcal{F}_\mu^\nu), d_{2,0})$  (with  $\beta = 1, D = 0$  in (4.6) using proposition 2.11). Note also that the approximation property in the definition implies  $\nabla_x g(\tau_{x,u}) \in L^2(W^*(x, u))$ .

The application of our Theorem thus gives a unique solution  $X_t(X_0)$  on  $[0, \infty[$  solving

$$X_t(X_0) = X_0 - \int_0^t \nabla G(X_s(X_0)) ds - \int_0^t X_s(X_0) ds + S_t.$$

Considering another solution starting at  $Y_0$ , one obtains :

$$\begin{aligned}
\|X_t(X_0) - X_t(Y_0)\|_2^2 &= \|X_0 - Y_0\|_2^2 - \int_0^t \langle \nabla G(X_s(X_0)) - \nabla G(X_s(Y_0)), X_s(X_0) - X_s(Y_0) \rangle ds \\
&\quad - \int_0^t \langle X_s(X_0) - X_s(Y_0), X_s(X_0) - X_s(Y_0) \rangle ds \\
&\leq \|X_0 - Y_0\|_2^2 - \int_0^t \langle X_s(X_0) - X_s(Y_0), X_s(X_0) - X_s(Y_0) \rangle ds,
\end{aligned}$$

where the last inequality comes from  $g(\tau_{X,u})$  convex. Applying Gronwall's lemma, one gets the stated exponential decay. But [D10b, Theorem 28] since  $\tau_{X^\omega,u}$  has finite Fisher information, there is a stationary solution to the same equation. But by the uniqueness of our solution in Theorem 4.3, the solution must be this same stationary process. But exponential decay implies that the laws  $\tau_{X_t(X^\omega)}$  and  $\tau_{X_t(X^{\omega'})}$  are arbitrarily close for  $t \rightarrow \infty$  and since they are equal to  $\tau_{X^\omega}$  and  $\tau_{X^{\omega'}}$  by stationarity, one deduces that  $X^\omega$  have the same law for any ultrafilter. Similarly,  $(SD_g)$  has a unique solution and the exponential decay implies a stationary state for the SDE is unique too.

**Step 4 :** Conclusion on the limit of  $E_{\mu_{g,N}} \circ \tau_\cdot$ .

The law  $E_{\mu_{g,N}} \circ \tau_{\cdot, \Upsilon_N}$  is close to  $E_{\mu_{g,N}} \circ \tau_{B^N, \Upsilon_N}$  for  $N$  large enough and this second law lies in the compact set  $S_C^m * \mathcal{F}_\mu^\nu$  (for the weak-\* topology induced by  $d_{2,0}$ ) and from the result on ultrafilter limits the sequence has a unique limit point there (any such limit point being a  $\tau_{X^\omega,u}$ ). We thus deduce by compactness the expected convergence.  $\square$

## 5. MINIMIZATION IN BOUÉ-DUPUIS-ÜSTÜNEL FORMULA FOR HERMITIAN BROWNIAN MOTION

The key for our large deviation estimate is to use the results from [Us14] on the minimization problem in Boué-Dupuis-Üstünel formula (applied to hermitian brownian motion) for specifically nice functionals coming from  $\mathcal{E}_\alpha(\mathbb{R}^{dk})$ , and deduce an equivalent minimization problem better suited to take the large  $N$  matrix limit. Note that we make a sign correction in the SDE agreeing with [L, Theorems 2,4], which differs from [Us14].

**Theorem 5.1** (Theorem 11 in [Us14]). *Assume that  $f \in L^0(\gamma)$  is 1-convex and that  $f^- = \max(-f, 0)$  is exponentially integrable, i.e.,  $E[\exp cf^-] < \infty$  for some  $c > 1$ . Then there is a unique  $u \in L_a^2(\gamma, H)$  reaching the infimum appearing in the definition of tame functionals, provided that  $E[f \circ (I_W + \xi)] < \infty$  for at least one  $\xi \in L_a^2(\gamma, H)$ . Moreover, if  $f \in L^{1+\epsilon}(\gamma)$  for  $\epsilon = \frac{1}{c-1} > 0$ , then  $f$  is a tamed functional and if*

$$\dot{v}_t = - \frac{E[D_t e^{-f} | \mathcal{F}_t]}{E[e^{-f} | \mathcal{F}_t]}$$

(where formally  $E[D_t e^{-f} | \mathcal{F}_t] = [\hat{\pi} \nabla(e^{-f})]_t$  as in subsection 2.4), then  $U_t = B_t + u_t$  is the unique strong solution of the following stochastic differential equation:

$$dU_t = -\dot{v}_t \circ U dt + dB_t.$$

As a consequence, for this solution  $U^f := U$ , we have:

$$-\log \left( \int_{\mathbb{W}} e^{-f} d\gamma \right) = \mathbf{E} \left( f(U) + \frac{1}{2} \int_0^1 \|\dot{v}_t \circ U\|_2^2 dt \right).$$

We deduce from that crucial description of the minimizer in the convex case the result we need :

**Corollary 5.2.** *Let  $t_0 = 0 < t_1 < \dots < t_k \in [0, 1]$ ,  $g \in \mathcal{E}_2(\mathbb{R}^{dk})$ , and*

$$f = g \circ J_{t_1, \dots, t_k} : \mathbb{W} \rightarrow \mathbb{R}.$$

*Let  $b(t, \cdot)$  defined in corollary 4.2 and  $X(t)$  the unique strong solution starting at  $X_0 = 0$  defined there of*

$$X(t) = X_0 + \int_0^t b(s, X(s)) ds + B_t.$$

*Then, we have the formula:*

$$-\log \left( \int_{\mathbb{W}} e^{-f} d\gamma \right) = \mathbf{E} \left( f(X) + \frac{1}{2} \int_0^1 \|b(t, X_t)\|_2^2 dt \right).$$

*Proof.* Since  $g$  is convex,  $f$  is 0-convex thus 1-convex and since  $g$  is bounded from below,  $f^-$  is exponentially integrable. If  $\xi = 0$ , the subquadratic bound of  $g$  implies  $E[f \circ (I_W + \xi)] < \infty$  and  $f \in L^{1+\epsilon}(\gamma)$  for any  $\epsilon > 0$ . Thus, the previous theorem applies,  $f$  is a tame functional. It essentially remains to identify  $X(t) = U_t^f$  with the solution considered in the theorem.

For, recall that by definition, we have the inductive definition (using the solution  $X_{t_k}(\omega)$  for small times with  $k \leq i$ ) for  $t_i < t \leq t_{i+1}$  :

$$b_j(t, x, \omega) = -\frac{\partial}{\partial x^{(j)}} h_t(X_{t_1}(\omega), \dots, X_{t_i}(\omega), x),$$

and we have from the proof of proposition 3.5:

$$h_t(x_1, \dots, x_i, x) = -\log \left( \int_{\mathbb{W}_{[t,1]}} e^{-g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} d\gamma_{[t,1]}(\nu) \right).$$

From the differentiability of  $g \in \mathcal{E}_2(\mathbb{R}^{dk})$  and the lipschitzness bound (3.3) implying the boundedness of derivatives. We have the bound :

$$\begin{aligned} & \left| \frac{\partial}{\partial x^{(j)}} e^{-g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} \right| \\ &= e^{-g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} \left| \frac{\partial}{\partial x^{(j)}} (g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})) \right| \\ &\leq e^{\sup(-g)} \sum_{l=i+1}^k \left( 2C \|x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k}\|_2 + D\sqrt{d} \right). \end{aligned}$$

which is an integrable dominating function on compact sets for  $x$ , so that one can compute the derivative by derivation of integral depending on parameters under Lebesgue domination condition and using that

$$\left( \frac{\partial}{\partial x^{(j)}} e^{-g(B_{t_1}, \dots, B_{t_i}, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} \right)_{x=B_t} = D_t(e^{-g \circ J_{t_1, \dots, t_k}})_{\omega=B},$$

one gets the identity (taking  $D_t$  with respect to the process  $\tilde{\omega} + \nu$  in path space):

$$(5.1) \quad b_j(t, X_t(\omega), \omega) = \frac{\left( \int_{\mathbb{W}_{[t,1]}} D_t e^{-g(\tilde{\omega}_{t_1}, \dots, \tilde{\omega}_{t_i}(\omega), \tilde{\omega}_t + \nu_{t_{i+1}}, \dots, \tilde{\omega}_t + \nu_{t_k})} d\gamma_{[t,1]}(\nu) \right)_{\tilde{\omega}=X_t(\omega)}}{\left( \int_{\mathbb{W}_{[t,1]}} e^{-g(X_{t_1}(\omega), \dots, X_{t_i}(\omega), x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} d\gamma_{[t,1]}(\nu) \right)},$$

which is rewritten  $b(t, X_t(\omega), \omega) = -\dot{v}_t \circ X$  and thus implies that the  $U$  of the theorem, satisfies the same equation as  $X$ , thus,  $U^f = X$  and from the equality of the drift, one obtains the stated equality of integrals reaching the infimum.  $\square$

We finally apply our results to hermitian brownian motion of section 2.4. We start by a lemma relating our various classes of functionals.

**Lemma 5.3.** *Let  $p \in [2, \infty]$ ,  $d = N^2 m$  and  $G \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$ . Seeing  $\mathbb{R}^{dk} = ((M_N(\mathbb{C})_{sa})^m)^k$ , for  $x = (H_1, \dots, H_k) \in \mathbb{R}^{dk}$ ,  $H_i \in (M_N(\mathbb{C})_{sa})^m$ , and  $\Upsilon_N \in \mathcal{U}((M_N(\mathbb{C}))^{\mu\nu})$ , we let  $\tau_x, \Upsilon_N \in \mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu)$  the corresponding state. Then  $g : x \mapsto N^2 G(\tau_{x/\sqrt{N}}, \Upsilon_N) \in \mathcal{E}_\alpha(\mathbb{R}^{dk})$  for any  $\alpha \in [1, 2]$  with constants independent of  $N, \Upsilon_N$  in (3.2), (3.3) and (3.4). In case  $\alpha = 2$ , one can even take  $D_2 = 0$ , and in this case the same result holds for  $G \in \mathcal{E}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$ .*

*Proof.* The case  $G \in \mathcal{E}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  is immediate by change of variable: (3.2) comes from (2.1), (3.3) comes from (2.2) and (3.4) from (2.4).

The convexity of  $g$  comes from the universal convexity of  $G$  and the continuity from the one of  $G$  once noted that  $x \mapsto \tau_{x/\sqrt{N}}, \Upsilon_N$  is continuous with value  $(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$ .

From the subquadratic behaviour in the definition of  $\mathcal{E}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , one gets:

$$g(x) = N^2 G(\tau_{x/\sqrt{N}}) \leq C(N^2 + N^2 \sum_{j=1}^k \sum_{l=1}^m \frac{1}{N^2} Tr(((4i \frac{u_j^l + 1}{u_j^l - 1})^* (4i \frac{u_j^l + 1}{u_j^l - 1}))) = C(N^2 + \|x\|_2^2),$$

as expected with a constant  $C$  independent of  $N$ . Moreover, for  $G(\tau) = D + \left( \sum_{i=1, \dots, l} (g_i(\tau))^p \right)^{1/p}$  as in the definition of  $\mathcal{E}_{reg,p}(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$  one can differentiate in  $x = (x_{i,j}^k)_{i,j=1, \dots, N; k=1, \dots, m}$ <sup>1</sup> since the values of  $g_i(\tau) \geq 1$  are not close of the point where the root is not differentiable:

$$\begin{aligned} & \left\| \left( \frac{\partial}{\partial x_{i,j}^k} g(x) \right)_{i,j,k} \right\| \\ &= N^2 \frac{1}{p} \left( \sum_{l=1, \dots, l} \left( g_l(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^p \right)^{(1-p)/p} \left\| \sum_{l=1, \dots, l} p \left( g_l(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^{p-1} \frac{\partial}{\partial x_{i,j}^k} g_l(\tau_{x/\sqrt{N}}, \Upsilon_N) \right\| \\ &\leq N^2 \left( \sum_{l=1, \dots, l} \left\| \left( \frac{\partial}{\partial x_{i,j}^k} g_l(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)_{i,j,k} \right\|^p \right)^{1/p} \end{aligned}$$

where we applied Hölder inequality of exponents  $p, q = p/(p-1)$ . Then, recall that  $g_l(\tau) = D_l + C_l \sum_{j=1}^k \sum_{l=1}^m \tau((4i \frac{u_j^l + 1}{u_j^l - 1})^* (4i \frac{u_j^l + 1}{u_j^l - 1})) + \Re(\lambda_l \tau((u_{j_1}^{i_1})^{\epsilon_1} \dots (u_{j_{m_l}}^{i_{m_l}})^{\epsilon_{m_l}}))$  so that as in [BCG, section 3.2], one gets:

$$\frac{\partial}{\partial x_{i,j}^k} g_l(\tau_{x/\sqrt{N}}, \Upsilon_N) = \frac{2C_l}{N^2} x_{i,j}^k + \Re \left( \frac{2C_l \lambda_l}{N \sqrt{N}} \lambda_{i,j} \left( \sum_{l=1}^{m_l} ((u_{j_l}^{i_l})^{\epsilon_l} - 1) \dots (u_{j_{m_l}}^{i_{m_l}})^{\epsilon_{m_l}} (u_{j_1}^{i_1})^{\epsilon_1} \dots ((u_{j_l}^{i_l})^{\epsilon_l} - 1) \frac{\epsilon_l}{8\sqrt{-1}} 1_{k=j_l} \right) \right)$$

so that one can estimate  $\left\| \left( \frac{\partial}{\partial x_{i,j}^k} g_l(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)_{i,j,k} \right\|^2 \leq \frac{8C_l^2}{N^4} \|x\|^2 + \frac{8C_l^2 |\lambda_l|^2 m_l^2 N m}{4N^3}$  and get finally :

$$\left\| \left( \frac{\partial}{\partial x_{i,j}^k} g(x) \right)_{i,j,k} \right\| \leq \max_l (8C_l^2 \|x\|^2 + 2C_l^2 |\lambda_l|^2 m_l^2 N^2 m)^{1/2} l^{1/p}$$

Thus applying the fundamental theorem of calculus, one gets (3.3) with  $d = N^2 m$  and constants  $C, D$  independent on  $N$ .

<sup>1</sup>so that the  $M_n(\mathbb{C})$  matrix entries are  $x_{i,i}^k$  and  $z_{i,j}^k = x_{i,j}^k + \sqrt{-1} x_{j,i}^k$  for  $i < j$  in index  $(i, j)$  and deduced by hermitianity in index  $(j, i)$ , and we call

$$\lambda_{i,j}(x^k) = x_{i,j}^k = \sqrt{(-1)^{1_{j < i}}} (z_{i,j}^k + (-1)^{1_{j < i}} z_{j,i}^k) / 2$$

the linear application realizing this choice of matrix entries for  $i \neq j$  and  $\lambda_{i,i}$  the matrix entry also defined on non-hermitian matrices if necessary in that way.

To get the second order bound, we are going to compute the second order derivative.

$$\begin{aligned}
& \left| \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \left( \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g(x) \right) \right| \leq N^2 \frac{1}{p} \left( \sum_{\iota=1,\dots,l} \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^p \right)^{1/p-1} \\
& \times \left[ \left| \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \sum_{\iota=1,\dots,l} p(p-1) \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^{p-2} \frac{\partial}{\partial x_{i,j}^k} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \frac{\partial}{\partial x_{I,J}^K} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right| \right. \\
& + \left| \frac{1}{p} - 1 \right| \left( \sum_{\iota=1,\dots,l} \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^p \right)^{-1} \left\| \sum_{\iota=1,\dots,l} p \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^{p-1} \frac{\partial}{\partial x_{i,j}^k} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right\|^2 \|\lambda\|^2 \\
& + \left. \left| \sum_{\iota=1,\dots,l} p \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^{p-1} \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right| \right]
\end{aligned}$$

Thus applying Holder's inequality as before (for  $p \geq 2$  for the first inequality which yields the same type of terms as in the second line and are thus gathered) :

$$\begin{aligned}
& \left| \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \left( \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g(x) \right) \right| \\
& \leq 2N^2 p \left( \sum_{\iota=1,\dots,l} \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^p \right)^{-1/p} \left( \sum_{\iota=1,\dots,l} \left\| \left( \frac{\partial}{\partial x_{i,j}^k} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)_{i,j,k} \right\|^p \right)^{2/p} \|\lambda\|^2 \\
& + N^2 \left( \sum_{\iota=1,\dots,l} \left| \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right| \right)^{1/p}
\end{aligned}$$

Thus, we have to compute the second derivative

$$\begin{aligned}
& \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) = \frac{2C_\iota}{N^2} 1_{i=I} 1_{j=J} 1_{k=K} + \Re \left( \frac{2C_\iota \lambda_\iota}{N^2} \sum_{l=1}^{m_\iota} \sum_{L \in [1, m_\iota] - \{l\}} \frac{-\epsilon_L \epsilon_l}{8^2} \right. \\
& \left. \lambda_{I,J} \left( \lambda_{i,j} \left( \left( (u_{j_l}^{i_l})^{\epsilon_l} - 1 \right) \dots \left( (u_{j_L}^{i_L})^{\epsilon_L} - 1 \right) \right)_{i'I} \left[ \left( (u_{j_L}^{i_L})^{\epsilon_L} - 1 \right) 1_{K=j_L} \dots \left( (u_{j_l}^{i_l})^{\epsilon_l} - 1 \right) 1_{k=j_l} \right]_{J'j'} \right)_{i'j'} \right)_{I'J'} \right)
\end{aligned}$$

so that  $\left| \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right| \leq \frac{2C_\iota}{N^2} \|\lambda\|^2 + \frac{cC_\iota \lambda_\iota m_\iota^2}{N^2} \|\lambda\|^2$  for some constant  $c$  not depending on  $\iota, N$ .

Finally, to get a better estimate, one needs a lower bound on  $\left( \sum_{\iota=1,\dots,l} \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^p \right) \geq l$  but also  $g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \geq C_i \left\| \frac{x}{\sqrt{N}} \right\|^2 - E_i$  so that one deduces for  $\left\| \frac{x}{\sqrt{N}} \right\|^2 \geq 2E_i/C_i + 1/C_i$ , one deduces  $g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \geq C_i \left\| \frac{x}{\sqrt{N}} \right\|^2/2 + 1/2$  and for  $C_i \left\| \frac{x}{\sqrt{N}} \right\|^2/2 \leq E_i + 1/2$  one can fix  $q \geq 1$  such that  $(E_i + 1/2) \leq q/2$  so that  $C_i \left\| \frac{x}{\sqrt{N}} \right\|^2/2q \leq 1/2$  and  $g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \geq 1 \geq 1/2 + C_i \left\| \frac{x}{\sqrt{N}} \right\|^2/2q$  and in any case :

$$\left( \sum_{\iota=1,\dots,l} \left( g_\iota(\tau_{x/\sqrt{N}}, \Upsilon_N) \right)^p \right) \geq l - 1 + \left( \frac{1}{2} + \frac{C_1}{2q} \left\| \frac{x}{\sqrt{N}} \right\|_2^2 \right)^p.$$

Thus gathering all our inequalities, one gets for a constant  $c_1$  independent of  $N, x$ :

$$\left| \sum_{i,j,k,I,J,K} \lambda_{i,j,k} \lambda_{I,J,K} \left( \frac{\partial^2}{\partial x_{i,j}^k \partial x_{I,J}^K} g(x) \right) \right| \leq c_1 \|\lambda\|^2$$

Using the fundamental theorem of calculus, one deduces a bound on second order difference quotients (using in the next-to-last line convexity of the squared euclidian norm):

$$\begin{aligned} |g(x+y) + g(x-y) - 2g(x)| &= \left| \int_0^1 d\lambda \int_0^1 d\mu d^2g(x - \lambda y + 2\lambda\mu y)(y, 2\lambda y) \right| \\ &\leq \int_0^1 d\lambda \int_0^1 d\mu c_1 \lambda \|y\|^2 \leq c_1 \|y\|^2 \end{aligned}$$

Combining this with the Lipschitz bound previously obtained of the form

$$|g(x+y) + g(x-y) - 2g(x)| \leq N(c_2 \frac{\|x\| + \|x-y\| + \|x+y\|}{N} + c_3) \|y\|$$

one gets (3.4) with constants independent of  $N, d = N^2m$  in taking the power  $\alpha - 1$  of our second order bound and multiplied by the power  $(2 - \alpha)$  of our lipschitz bound.  $\square$

We can now gather our results in the matricial case, as needed to be applied to prove the Laplace principle in the next section. Recall that the natural euclidean norm for  $x \in ((M_N(\mathbb{C}))_{sa})^m$  is given by the following notation we will use in our next result:

$$\|x\|_2^2 = \sum_{k=1}^m \frac{1}{N} \text{Tr}(x_k^* x_k).$$

It will also be convenient to use the matricial cyclic gradient for functions  $h : (M_N(\mathbb{C}))_{sa}^{mi} \rightarrow \mathbb{R}$ , for  $x_i \in ((M_N(\mathbb{C}))_{sa})^m$ ,  $I < J \in \llbracket 1, N \rrbracket$ ,  $k = 1, \dots, m$ ,  $j = 1, \dots, i$  (with the notation for coordinates of footnote <sup>1</sup>)

$$\begin{aligned} (\mathcal{D}_j^k h)_{II} &:= \left( \frac{\partial}{\partial (x_j)_{II}^k} h \right) (x_1, \dots, x_i), \\ (\mathcal{D}_j^k h)_{IJ} &:= \left( \frac{\partial}{\partial (x_j)_{IJ}^k} h + \sqrt{-1} \frac{\partial}{\partial (x_j)_{JI}^k} h \right) (x_1, \dots, x_i), \\ (\mathcal{D}_j^k h)_{JI} &:= \left( \frac{\partial}{\partial (x_j)_{JI}^k} h - \sqrt{-1} \frac{\partial}{\partial (x_j)_{IJ}^k} h \right) (x_1, \dots, x_i), \end{aligned}$$

**Theorem 5.4.** *Let  $p \in [2, \infty[$ ,  $d = N^2m$   $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))$  and  $G \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \cup \mathcal{E}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  so that  $g_N(x) = N^2 G(\tau_{x/\sqrt{N}, \Upsilon_N}) \in \mathcal{E}_2(\mathbb{R}^{dk})$  as in the previous lemma. Let  $t_0 = 0 < t_1 < \dots < t_k \in [0, 1]$  and  $f_N = g_N \circ J_{t_1, \dots, t_k} : \mathbb{W}_{sa, N} \rightarrow \mathbb{R}$ .*

*Let, for  $x_i, x \in ((M_N(\mathbb{C}))_{sa})^m$  and  $t \in [0, 1]$ :*

$$h_t^{N, \Upsilon_N}(x_1, \dots, x_i, x) = -\log \left( \int_{\mathbb{W}_{sa, N, [t, 1]}} e^{-g_N(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} d\gamma_{[t, 1]}(\nu) \right)$$

*and for  $\omega \in \mathbb{W}_{sa, N}$ ,  $k = 1, \dots, m$  define inductively on  $i$  for  $t_i < t \leq t_{i+1}$ :*

$$(b_k^{G, N, \Upsilon_N}(t, x, \omega)) := -\frac{1}{\sqrt{N}} \left( \mathcal{D}_{i+1}^k h_t^{N, \Upsilon_N} \right) (\sqrt{N} X_{t_1}^{G, N, \Upsilon_N}(\omega), \dots, \sqrt{N} X_{t_i}^{G, N, \Upsilon_N}(\omega), \sqrt{N} x),$$

*so that there is a unique (strong) solution to the  $m$ -tuple of matrix-valued SDE driven by  $H_t$  brownian motion of law  $\gamma_{sa, N, m}$ :*

$$X^{G, N, \Upsilon_N}(t) = \int_0^t b^{G, N, \Upsilon_N}(s, X^{G, N, \Upsilon_N}(s)) ds + H_t^N.$$

*Then, we have the formula :*

$$-\frac{1}{N^2} \log \left( \int_{\mathbb{W}_{sa, N}} e^{-g_N(\nu_{t_1}, \dots, \nu_{t_k})} d\gamma(\nu) \right) = \mathbf{E} \left( G(\tau_{X^{G, N, \Upsilon_N}, \Upsilon_N}) + \frac{1}{2} \int_0^1 \|b^{G, N, \Upsilon_N}(t, X^{G, N, \Upsilon_N}(t))\|_2^2 dt \right).$$

*Moreover we have for some  $C_4 = C_4(G)$  independent of  $N$  and for all  $t, s \in ]t_i, t_{i+1}]$ :*

$$E(\|b^{G, N, \Upsilon_N}(t, X_t^{G, N, \Upsilon_N}) - b^{G, N, \Upsilon_N}(s, X_s^{G, N, \Upsilon_N})\|_2^2) \leq C_4 \sqrt[4]{|t - s|},$$



$$|h_t^{N, \Upsilon_N}(x_1, \dots, x_i, x) - h_t^{N, \Upsilon'_N}(x_1, \dots, x_i, x)| \leq C(G) \|\Upsilon_N - \Upsilon'_N\|_2.$$

Finally, the law of  $\sqrt{N}X^{G, N, \Upsilon_N}(t)$  on the pathspace  $\mathbb{W}_{sa, N}$  is exactly the Gibbs law:

$$\frac{e^{-g_N(\nu_{t_1}, \dots, \nu_{t_k})}}{\int d\gamma(\nu) e^{-g_N(\nu_{t_1}, \dots, \nu_{t_k})}} d\gamma(\nu).$$

*Proof.*  $\sqrt{N}X_s^{G, N, \Upsilon_N}$  is the solution from Corollary 4.2 with our  $g$  which satisfies the assumption by our previous lemma 5.3. The formula is then exactly the one given by Corollary 5.2. The first bound is also given in Corollary 4.2. The independence of  $N$  comes from the dimension independence of the constants given in lemma 5.3 that make all the corresponding constants in proposition 3.5 also independent of dimension. For the final statement, we can use for instance the well-known [Us14, Theorem 5] to get:

$$\begin{aligned} & \mathbf{E} \left( G(\tau_{X^{G, N, \Upsilon_N}, \Upsilon_N}) + \frac{1}{2} \int_0^1 \|b^{G, N, \Upsilon_N}(t, X^{G, N, \Upsilon_N}(t))\|_2^2 dt \right) \\ &= \inf \left\{ \int G(\tau, \Upsilon_N) d\mu + \frac{1}{N^2} \text{Ent}(\mu|\gamma), \mu \in \text{Prob}(\mathbb{W}_{sa, N}) \right\} \end{aligned}$$

where  $\text{Ent}(\mu|\gamma)$  is the usual relative entropy so that for instance [Us14, Theorem 2] gives that for  $\mu_{G, N, \Upsilon_N}$  the law of  $\sqrt{N}X^{G, N, \Upsilon_N}$

$$\text{Ent}(\mu_{G, N, \Upsilon_N}|\gamma) \leq \frac{N^2}{2} \int_0^1 \|b^{G, N, \Upsilon_N}(t, X^{G, N, \Upsilon_N}(t))\|_2^2 dt.$$

Thus the law  $\mu_{G, N, \Upsilon_N}$  reaches the infimum above and the uniqueness of the infimum and its form as a Gibbs state are given in the already quoted [Us14, Theorem 5].

Finally, the lipschitz bound in  $\Upsilon_N$  is obtained as in step 3 of the proof of proposition 3.5 but using (2.3) instead of (2.2).  $\square$

### 5.1. Alternative formula for the drift $b^{G, N, \Upsilon_N}$ .

**Proposition 5.5.** *Consider the setting of Corollary 5.2. Let  $t_i < t \leq t_{i+1}$  and  $(x_1, \dots, x_i, x) \in \mathbb{R}^{d(i+1)}$ .*

*We define  $b^t(s, y, \omega)$  for  $s \in [t, 1]$  and  $\omega \in \Omega$ , in considering the case  $\max(t_I, t) < s \leq t_{I+1}, I \geq i$  by :*

$$b_j^t(s, y, \omega) := -\frac{\partial}{\partial y^{(j)}} h_s(x_1, \dots, x_i, X_{t_{i+1}}^t(\omega), \dots, X_{t_I}^t(\omega), y),$$

*and  $b^t(s, y, \omega) = 0$  if  $s > t_k$ . We define simultaneously  $X^t(s) = X^t(s, x)$  the unique strong solution starting at  $X_t^t = x$  defined in Corollary 4.2 of*

$$X^t(s) = X_t^t + \int_t^s b(u, X^t(u)) du + B_s - B_t.$$

*Also define for convenience  $X_{t_I}^t = x_I$ ,  $x_0 = 0$  for  $I \leq i$  (and say interpolate linearly values on  $[0, t]$ ). Then, we have the formulas:*

$$h_t(x_1, \dots, x_i, x) = \mathbf{E} \left( f(X^t) + \frac{1}{2} \int_t^1 \|b(s, X^t(s))\|_2^2 ds \right),$$

$$\frac{\partial}{\partial x^{(j)}} h_t(x_1, \dots, x_i, x) = \mathbf{E} \left( \left[ \frac{\partial}{\partial y^{(j)}} g(x_1, \dots, x_i, y + (X^t(t_{i+1}, x) - x), \dots, y + (X^t(t_k, x) - x)) \right]_{y=x} \right).$$

*Proof.* Consider the minimization problem in (3.6). Since for each  $X = (x_1, \dots, x_i, x)$  fixed,  $g_X(y_{i+1}, \dots, y_k) = g(x_1, \dots, x_i, x + y_{i+1}, \dots, x + y_k)$  defines a function  $g_X \in \mathcal{E}_2(\mathbb{R}^{d(k-i)})$ , we can apply Corollary 5.2 to it. The first formula for  $h_t$  is then exactly the result of this corollary. Arguing as in the proof of Theorem 5.4, one knows that the law of  $(X_s - x)_{s \in [t, 1]}$  is

$$\frac{e^{-g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} d\gamma_{[t, 1]}(\nu)}{\int_{\mathbb{W}_{[t, 1]}} e^{-g(x_1, \dots, x_i, x + \nu_{t_{i+1}}, \dots, x + \nu_{t_k})} d\gamma_{[t, 1]}(\nu)}.$$

But this measures appears in (5.1) (and its variant with  $X_t$  replaced by a generic point  $(x_1, \dots, x_i, x)$ ) which we can interpret as the second expected formula.  $\square$

For convenience, we state separately the obvious application in the matricial case.

**Corollary 5.6.** *Fix the setting of Theorem 5.4. Let  $t_i < t \leq t_{i+1}$  and  $(x_1, \dots, x_i, x) \in ((M_N(\mathbb{C}))_{sa})^{m(i+1)}$ .*

*We define  $b^{G,N,\Upsilon_N,t}(s, y, \omega)$  for  $s \in [t, 1]$  and  $\omega \in \Omega$ , in considering the case  $\max(t_I, t) < s \leq t_{I+1}$ ,  $I \geq i$  by :*

$$b_k^{G,N,\Upsilon_N,t}(s, y, \omega) := -\frac{1}{\sqrt{N}} \left( \mathcal{D}_{i+1}^k h_s^{N,\Upsilon_N} \right) (\sqrt{N}x_1, \dots, \sqrt{N}x_i, \sqrt{N}X_{t_{i+1}}^{G,N,\Upsilon_N}(\omega), \dots, \sqrt{N}X_{t_I}^{G,N,\Upsilon_N}(\omega), \sqrt{N}y),$$

*and  $b^{G,N,\Upsilon_N,t}(s, y, \omega) = 0$  if  $s > t_k$ . We define simultaneously  $X^{G,N,\Upsilon_N,t}(s) = X^{G,N,\Upsilon_N,t}(s, x)$  the unique strong solution starting at  $X^{G,N,\Upsilon_N,t}(t) = x$  defined in Corollary 4.2 driven by  $H_t^N$  brownian motion of law  $\gamma_{sa,N,m}$  of*

$$X^{G,N,\Upsilon_N,t}(s) = X^{G,N,\Upsilon_N,t}(t) + \int_t^s b^{G,N,\Upsilon_N,t}(u, X^{G,N,\Upsilon_N,t}(u)) du + H_s^N - H_t^N.$$

*Also define for convenience  $X^{G,N,\Upsilon_N,t}(t_I) = x_I$ ,  $x_0 = 0$  for  $I \leq i$  (and say interpolate linearly values on  $[0, t]$ ). Then, we have the formulas :*

$$\begin{aligned} & \frac{1}{N^2} h_t^{N,\Upsilon_N}(\sqrt{N}x_1, \dots, \sqrt{N}x_i, \sqrt{N}x) \\ &= \mathbf{E} \left( G(\tau_{X^{G,N,\Upsilon_N,t}, \Upsilon_N}) + \frac{1}{2} \int_s^1 \|b^{G,N,\Upsilon_N,t}(u, X^{G,N,\Upsilon_N,t}(u))\|_2^2 du \right), \\ & (\mathcal{D}_{i+1}^l h_t^{N,\Upsilon_N})(\sqrt{N}x_1, \dots, \sqrt{N}x_i, \sqrt{N}x) \\ &= \mathbf{E} \left( \sum_{j=i+1}^k \mathcal{D}_j^l g_N(\sqrt{N}x_1, \dots, \sqrt{N}x_i, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_{i+1}, x), \dots, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_k, x)) \right). \end{aligned}$$

## 6. LAPLACE PRINCIPAL FOR HERMITIAN BROWNIAN MOTION

We first want to define the candidate for the limiting function in the Laplace principle for  $f \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ .

We fix  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))$  (deterministic). Assume finally that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\Upsilon \in (\mathcal{T}(\mathcal{F}_\mu^\nu), d)$ . Recall that we aim at proving a Laplace principle for  $\widehat{\sigma}_\Upsilon^N$  the joint law of  $\Upsilon_N$  and an hermitian brownian motion.

Recall  $\gamma_{sa,N,m} = \gamma_N$  is the law of the hermitian brownian motion.

We can consider the von Neumann algebra  $M_N(L^\infty(\mathbb{W}_{sa,N}, \gamma_N))$  in which lives random matrix processes over this probability space. This is a finite non-commutative probability space with trace  $\tau_{\gamma_N} = E_{\gamma_N} \circ \frac{1}{N} \text{Tr}$ . We recall we use  $\beta\mathbb{N} - \mathbb{N}$  the set of non-principal ultrafilters  $\omega \in \beta\mathbb{N} - \mathbb{N}$  on  $\mathbb{N}$ . One considers the tracial ultraproduct  $\mathcal{M}_P^\omega = (M_N(L^\infty(\mathbb{W}_{sa,N}, \gamma_N)), \tau_{\gamma_N})^\omega$ . Of course there is a natural filtration  $\mathcal{M}_{P,s}^\omega = (M_N(L^\infty(X_t, t \leq s)), \tau_{\gamma_N})^\omega$ , where  $X_t$  is the coordinate process of our hermitian process and  $L^\infty(X_t, t \leq s)$  is the generated commutative von Neumann algebra. We will also need  $\mathcal{L}_P^\omega = [L^2(M_N(L^\infty(\mathbb{W}_{sa,N}, \gamma_N)), \tau_{\gamma_N})]^\omega$  and  $\mathcal{L}_{P,s}^\omega = [L^2(M_N(L^\infty(X_t, t \leq s)), \tau_{\gamma_N})]^\omega$ .

Let  $M = W^*(v, S_t^i, i = 1, \dots, m, t \in [0, 1])$  the von Neumann algebra of a free brownian motion free from  $v$  of law  $\mu_\Upsilon$  and  $M_s = W^*(u, S_t^i, i = 1, \dots, m, t \in [0, s])$  its canonical filtration. Fix a sequence of times  $\mathbf{t} = (t_0 = 0 < t_1 < \dots < t_k \leq t_{k+1} = 1)$  We consider a set of adapted paths which are sufficiently Hölder continuous except at a sequence of times

$$\mathcal{P}_{ad,1/8,\mathbf{t}} = \{u \in L_{ad}^\infty([0, 1], L^2(M)^m) : \exists C > 0, \forall (s, t) \in ]t_i, t_{i+1}]^2, \|u_t - u_s\|_2 \leq C(t - s)^{1/8}\}$$

$$\mathcal{P}_{ad,1/8loc,\mathbf{t}} = \{u \in L_{ad}^\infty([0, 1], L^2(M)^m) :$$

$$\forall [a, b] \subset ]t_i, t_{i+1}[, \exists C > 0, \forall (s, t) \in ]a, b]^2, \|u_t - u_s\|_2 \leq C(t - s)^{1/8}\}$$

We consider another family of paths such that one requires more von Neumann algebraic regularity at this sequence of times:

$$\mathcal{P}_{ad,factor,\mathbf{t}} = \{u \in L_{ad}^\infty([0,1], L^2(M)^m) :$$

$$\forall i = 1, \dots, k, Y_{t_i} = S_{t_i} + \int_0^{t_i} u_s ds \in M^m \text{ and } W^*(v, Y_{t_1}, \dots, Y_{t_k}) \text{ is a factor} \}.$$

and the variant

$$\mathcal{P}_{ad,factor,\mathbf{t},b} = \{u \in L_{ad}^\infty([0,1], L^2(M)^m) :$$

$$\exists C \forall i = 1, \dots, k, \exists \tau_i < t_i < T_i, \forall t \in [\tau_i, T_i], Y_t = S_t + \int_0^t u_s ds \in M^m, \|Y_{t,l}\| \leq C$$

$$\text{and } W^*(v, Y_{t_1}, \dots, Y_{t_k}) \text{ is a factor} \}.$$

We are finally ready to define our candidate to be a limiting function for  $f \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  for  $p \in [1, \infty]$ . We even define it for any  $f \in C_{(k)}^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , bounded from below. For we call  $\mathbf{t}(f) = (t_1 < \dots < t_k)$  the minimal set of times such that  $f = g \circ (I_{t_1, \dots, t_k} * Id)$ .

$$\Lambda_v(f) := \inf \left\{ \frac{1}{2} \int_0^1 \|u_s\|^2 ds + f(\tau_{S+\int_0^\cdot u_s ds, v}) : u \in \mathcal{P}_{ad,1/8loc,\mathbf{t}(f)} \cap \mathcal{P}_{ad,factor,\mathbf{t}(f)} \right\}.$$

$$\Lambda_{b,v}(f) := \inf \left\{ \frac{1}{2} \int_0^1 \|u_s\|^2 ds + f(\tau_{S+\int_0^\cdot u_s ds, v}) : u \in \mathcal{P}_{ad,1/8,\mathbf{t}(f)} \cap \mathcal{P}_{ad,factor,\mathbf{t}(f),b} \right\}.$$

We give a name to the piecewise linear part

$$\mathcal{P}_{ad,1/8loc,\mathbf{t},pl} = \{u \in \mathcal{P}_{ad,1/8loc,\mathbf{t}} :$$

$$\exists s_0^1 = 0 < s_1^1 < \dots < s_n^1 < \dots < s_\omega^1 = s_0^2 < \dots < s_\omega^2 < \dots < s_\omega^k = s_0^{k+1} < s_1^{k+1} < \dots < s_l^{k+1}$$

$$\forall i, s_\omega^i = t_i, \forall i, j \geq 1, u_{s_i^j} \in L^2(M_{s_{i-1}^j})^m, u_{s_0^j} = u_{s_1^j}$$

$$\exists C, K > 0, \forall i > K, Y_{s_i^l} = S_{s_i^l} + \int_0^{s_i^l} u_s ds \in M^m, \|Y_{s_i^l, l}\| \leq C,$$

$$\forall s \in [s_i^j, s_{i+1}^j], u_s = \frac{s - s_i^j}{s_{i+1}^j - s_i^j} u_{s_{i+1}^j} + \frac{s_{i+1}^j - s}{s_{i+1}^j - s_i^j} u_{s_i^j}$$

We call  $\mathbf{s}(u)$  the minimal sequence of times appearing in the definition. Note that for  $u \in \mathcal{P}_{ad,1/8,\mathbf{t},pl}$ , we have:

$$(6.1) \quad \int_{s_i^l}^{s_{i+1}^l} u_s ds = (s_{i+1}^l - s_i^l) \frac{u_{s_{i+1}^l} + u_{s_i^l}}{2}.$$

We start by finding a first estimate for an alternative formula for  $\Lambda_{b,v}(f)$  that will be more convenient for the Laplace deviation upper bound since a piecewise linear functional depends locally on finitely many values and are easier to make converge in ultraproducts.

**Lemma 6.1.** *For any  $f \in C_{(k)}^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , bounded from below:*

$$\Lambda_v(f) \leq \inf \left\{ \frac{1}{2} \int_0^1 \|u_s\|^2 ds + f(\tau_{S+\int_0^\cdot u_s ds, v}) : u \in \mathcal{P}_{ad,1/8loc,\mathbf{t}(f),pl} \cap \mathcal{P}_{ad,factor,\mathbf{t}(f)} \right\} \leq \Lambda_{b,v}(f).$$

*Proof.* Clearly, we have the first  $\leq$  since  $\mathcal{P}_{ad,1/8loc,\mathbf{t}(f),pl} \subset \mathcal{P}_{ad,1/8loc,\mathbf{t}(f)}$ . For the second inequality, fix  $\mathbf{t}(f) = (t_0 = 0 < t_1 < \dots < t_k \leq t_{k+1} = 1)$ ,

Fix  $n \geq 2$  and define for  $k \geq l \geq 1$ ,

$$s_i^l = t_{l-1} + (t_l - t_{l-1}) \frac{i}{n}, \quad 0 \leq i \leq n-1$$

and

$$s_i^l = t_{l-1} + (t_l - t_{l-1}) \frac{n-1}{n} + (t_l - t_{l-1}) \frac{1}{n} \sum_{K=1}^{i-n+1} \frac{1}{2^K}, \quad i > n-1.$$

Fix  $v \in \mathcal{P}_{ad,1/4,\mathbf{t}(f)} \cap \mathcal{P}_{ad,factor,\mathbf{t}(f),b}$ , with Holder constant  $C$  and uniform bound  $\|v\|_\infty = \sup_{t \in [0,1]} \|v_t\|_2$ , and define  $u^{(n)} \in \mathcal{P}_{ad,1/4,\mathbf{t}(f),pl}$  as follows. First we take

$$u_{s_i^l}^{(n)} = v_{s_{i-1}^l}, \quad k \geq l \geq 1, \quad 1 \leq i \leq n-1$$

and  $u_{s_0^l}^{(n)} = v_{s_0^l}$  which is compatible with the measurability constraint  $i, j \geq 1$ ,  $u_{s_i^j} \in L^2(M_{s_{i-1}^j})^m$ . Thanks to the Hölder continuity of  $v$  this will guaranty a good uniform approximation. Of course we take a piecewise linear interpolation. We now want to guaranty properties near  $t_i$ , and for that we want

$$\int_{t_{l-1}}^{s_i^l} u_s^{(n)} ds = \int_{t_{l-1}}^{s_{i-1}^l} v_s ds, \quad k \geq l \geq 1, \quad i \in \llbracket n, \omega \rrbracket.$$

Since we expect a piecewise linear interpolation, one can use (6.1) to obtain :

$$\int_{t_{l-1}}^{s_i^l} u_s^{(n)} ds = \sum_{K=0}^{i-1} (s_{K+1}^l - s_K^l) \frac{u_{s_{K+1}^l}^{(n)} + u_{s_K^l}^{(n)}}{2}$$

This determines:

$$u_{s_n^l}^{(n)} = \frac{2}{(s_n^l - s_{n-1}^l)} \left( \int_{t_{l-1}}^{s_{n-1}^l} v_s ds - \sum_{K=0}^{n-2} \frac{(t_l - t_{l-1})}{n} \frac{u_{s_{K+1}^l}^{(n)} + u_{s_K^l}^{(n)}}{2} \right) - u_{s_{n-1}^l}^{(n)}$$

and then inductively  $u_{s_{i+1}^l}^{(n)}$  for  $i \geq n$ :

$$u_{s_{i+1}^l}^{(n)} = \frac{2}{(s_{i+1}^l - s_i^l)} \int_{s_{i-1}^l}^{s_i^l} v_s ds - u_{s_i^l}^{(n)}.$$

Note that those inductive definitions are compatible with the measurability constraint for  $i, j \geq 1$ ,  $u_{s_i^j} \in L^2(M_{s_{i-1}^j})^m$ .

Especially, we can bound inductively using the Hölder continuity of  $v$ , for  $i \geq n$ :

$$\begin{aligned} & \|u_{s_{i+1}^l}^{(n)} - v_{s_i^l}\|_2 \\ & \leq \frac{1}{(s_{i+1}^l - s_i^l)} \int_{s_{i-1}^l}^{s_i^l} \|v_s - v_{s_i^l}\|_2 ds + \frac{1}{(s_{i+1}^l - s_i^l)} \int_{s_{i-1}^l}^{s_i^l} \|v_s - v_{s_{i-1}^l}\|_2 ds + \|u_{s_i^l}^{(n)} - v_{s_{i-1}^l}\|_2 \\ & \leq 4C |s_{i-1}^l - s_i^l|^{1/8} + \|u_{s_i^l}^{(n)} - v_{s_{i-1}^l}\|_2 \leq 4C \sum_{I=n}^i |s_{I-1}^l - s_I^l|^{1/8} + \|u_{s_n^l}^{(n)} - v_{s_{n-1}^l}\|_2 \\ & \leq \frac{2C(t_l - t_{l-1})^{1/8}}{n^{1/8}} \sum_{I=n}^i \frac{1}{2^{(I-n+1)/8}} + \|u_{s_n^l}^{(n)} - v_{s_{n-1}^l}\|_2. \end{aligned}$$

One thus gets from the crude bound  $\|u_{s_n^l}^{(n)}\|_2 \leq (8n+1)\|v\|_\infty$  and by the converging geometric series in  $i$  above and convex combinations that  $u_s^{(n)}$  is bounded in  $L^2(M)$ .

From our construction we have:

$$\int_{t_{l-1}}^{t_l} u_s^{(n)} ds = \int_{t_{l-1}}^{t_l} v_s ds, \quad k \geq l \geq 1, \quad i \in \llbracket n, \omega \rrbracket,$$

and then by induction:

$$\int_0^{s_i^l} u_s^{(n)} ds = \int_0^{s_{i-1}^l} v_s ds, \quad k \geq l \geq 1, \quad i \in \llbracket n, \omega \rrbracket.$$

Thus for any  $n$ , one can deduce from  $Y_{t_i}(v) = Y_{t_i}(u^{(n)}) \in M^m$ , and the factoriality condition is thus also kept. We also deduce  $Y_{s_i^l}(v) = Y_{s_i^l}(u^{(n)}) \in M^m$ , for  $i$  large enough and the bounded constraint in  $M^m$  near  $t_l$  also follows from the one for  $v$ . We also have for the same reason

$f(S + \int_0^\cdot v_s ds) = f(S + \int_0^\cdot u_s^{(n)} ds)$  and from the bounds before and Hölder continuity of  $v$  one also easily prove

$$\frac{1}{2} \int_0^1 \|u_s^{(n)}\|^2 ds \rightarrow_{n \rightarrow \infty} \frac{1}{2} \int_0^1 \|v_s\|^2 ds.$$

It only remains to check the Hölder continuity condition for  $u^{(n)}$ . First if  $s_i^l$  is the smallest value above or the highest value below  $s$ , the linear interpolation implies  $\|u_s^{(n)} - u_{s_i^l}^{(n)}\| \leq M \|u^{(n)}\|_\infty |s - s_i^l|$  where  $M$  is finite as soon as extreme points remain within  $[a, b] \subset ]t_{l-1}, t_l[$  and we get Lipschitzness within intervals of  $[s_i^l, s_{i+1}^l]$  similarly. The Hölder continuity thus easily follows from the one with endpoints at times  $s_i^l, s_I^l$ . If these points are those with  $u^{(n)}$  coincide with values of  $v$  we are done, otherwise, we can take either  $i = n - 1, I \geq n$  or  $I > i \geq n$  but there are finitely many ratios with such endpoints in  $[a, b] \subset ]t_{l-1}, t_l[$ , thus there is necessarily even a Lipschitz bound.  $\square$

We now want to prove the Laplace principle we aim at getting. This is our main technical result:

**Theorem 6.2.** *Fix  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$  (deterministic). Assume that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\Upsilon \in (\mathcal{T}(\mathcal{F}_\mu^\nu), d)$ . We assume either  $m \geq 2$  or  $m = 1$  and  $W^*(\mu_\Upsilon)$  diffuse.*

*Let  $\gamma_{sa, N, m} = \gamma_N$  the law of hermitian  $N \times N$  brownian motion  $W_s^N \in (M_N(\mathbb{C}))^m$ , then, for any  $f \in \mathcal{E}_{reg, p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , for  $p \in [2, \infty]$  or  $f \in \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  the following limit exists and is given by our formula above :*

$$\lim_{N \rightarrow \infty} -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) = \Lambda_{b, v}(f).$$

The reader should note that what we will call lower bound (as in [BD]) corresponds to the usual large deviation upper bound.

*Proof. Step 1 :* Reduction of the case  $p = \infty$  to the limit for functionals in  $f \in \mathcal{E}_{reg, p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , for  $p \in [2, \infty[$

First given a functional in  $f \in \mathcal{E}_{reg, \infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , one can find  $g_i$  as in the definition and consider  $f_{(p)} \in \mathcal{E}_{reg, p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ ,

$$f_{(p)}(\bar{\tau}) = g_{(p)}(\bar{\tau} \circ (I_{t_1, \dots, t_k} * Id)) \quad \text{and} \quad g_{(p)}(\tau) = D + \left( \sum_{i=1, \dots, l} (g_i(\tau))^p \right)^{1/p}.$$

(the convexity condition on  $g_{(p)}$  is easily implied by those on  $g_i$ .) From the two next steps, we can assume the Laplace principle is satisfied for  $f_{(p)}$ .

Then by standard estimates between norms in finite dimension

$$g(\tau) \leq g_{(p)}(\tau) \leq l^{1/p} g(\tau).$$

As a consequence,

$$\begin{aligned} \limsup_{N \rightarrow \infty} -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) &\leq \limsup_{N \rightarrow \infty} -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f_{(p)}(\tau_W, \Upsilon_N)}) \\ &= \Lambda_{b, v}(f_{(p)}) \end{aligned}$$

and taking an infimum over  $p$  one easily get the upper bound  $\Lambda_{b, v}(f)$  since the set over which one takes the infimum does not depend on the argument of  $\Lambda$ , and  $f_{(p)} \rightarrow f$ . Conversely, we

obtain using the other inequality:

$$\begin{aligned} \liminf_{N \rightarrow \infty} \frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) &\geq \liminf_{N \rightarrow \infty} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f_{(p)}(\tau_W, \Upsilon_N)/l^{1/p}}) \\ &= \Lambda_{b,v}(f_{(p)}/l^{1/p}) \geq \Lambda_{b,v}(f/l^{1/p}) \end{aligned}$$

For a pair close enough to the infimum defining  $\Lambda_{b,v}(f/l^{1/p})$ , one can assume

$$f(\tau_{S+\int_0^\cdot u_s ds, v})/l^{1/p} \leq \frac{1}{2} \int_0^1 \|u_s\|^2 ds + \left( f(\tau_{S+\int_0^\cdot u_s ds, v}) \right) / l^{1/p} \leq f(\tau_{S,v})/l^{1/p} \leq f(\tau_{S,v}) = C(f)$$

the value at  $u = 0$ , which is a constant depending only on  $f$  since  $f$  depends only of the law and  $S$  is always a free brownian motion.

Thus one obtains the following concluding lower bound:

$$\Lambda_{b,v}(f/l^{1/p}) \geq \Lambda_{b,v}(f) - (l^{1/p} - 1)f(\tau_{S,v}) \xrightarrow{p \rightarrow \infty} \Lambda_{b,v}(f).$$

**Step 2 :** Lower bound for  $f \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \cup \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}), p \in [2, \infty[$ .

Consider an ultrafilter  $\omega \in \beta\mathbb{N} - \mathbb{N}$ . It suffices to show that :

$$(6.2) \quad \lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) \geq \Lambda_{b,v}(f).$$

We will use the tracial von Neumann algebra ultraproduct and the Hilbert space ultraproduct:

$$\mathcal{M}_P^\omega = (M_N(L^\infty(\mathbb{W}_{sa,N}, \gamma_N)), \tau_{\gamma_N})^\omega \subset \mathcal{L}_P^\omega = [L^2(M_{N_n}(L^\infty(\mathbb{W}_{sa,N}, \gamma_N)), \tau_{\gamma_N})]^\omega.$$

We can consider  $S_t = (W_t^N)^\omega \in \mathcal{L}_P^\omega$  and it is well known that actually,  $S_t \in \mathcal{M}_P^\omega$ . Indeed, from proposition 2.7, take  $C$  satisfying (2.8) so that if  $A_N = \{\|W_t^N\|_\infty \leq C\}$ , we have:

$$\limsup_{N \rightarrow \infty} \|W_t^N 1_{A_N} - W_t^N\|_2^2 = 0,$$

then  $S_t = (W_t^N 1_{A_N})^\omega \in \mathcal{L}_P^\omega$  and since by construction  $\|W_t^N 1_{A_N}\|_\infty \leq C$  one deduces that  $S_t \in \mathcal{M}_P^\omega$ , as expected. Of course we have  $v = (\Upsilon_N)^\omega \in (\mathcal{M}_{P,0}^\omega)^{\mu^\omega}$  which has consistently the law  $\mu_\Upsilon$ .

Unfortunately  $S_t$  is NOT a free brownian motion adapted to the canonical filtration  $\mathcal{M}_{P,s}^\omega$  (which is not a factor so that the covariance map of the process is a centered valued conditional expectation and not the trace). However,  $S_t$  is a free brownian motion adapted to its own filtration  $M_s = W^*(v, S_t, t \leq s) \subset \mathcal{M}_{P,s}^\omega$  (this is for instance a standard freeness result between GUE and constant matrices or one can use Theorem 2.20 and the concentration result proposition 2.7 for the computation of the covariance).

Note that we thus have a canonical (adapted) embedding  $I : M \subset \mathcal{M}_P^\omega$  which extends to  $I : L^2(M) \subset \mathcal{L}_P^\omega$ .

We will need to note that  $S_t$  and its stochastic integrals are still martingales adapted to  $\mathcal{L}_{P,s}^\omega$ , namely:

$$(6.3) \quad \forall U \in \mathcal{L}_{P,s}^\omega, \forall V \in L^2(M) \ominus L^2(M_s), \quad \langle U, V \rangle = 0.$$

From Clark-Ocone's formula (a slight extension of the one [BS] with extra initial conditions),  $V \in L^2(M) \ominus L^2(M_s)$  is a stochastic integral and thus can be approximated by a sum of terms of the form  $P \# (S_t - S_T)$ ,  $t \geq T \geq s$   $P \in \mathbb{C} \langle S_v, v \leq T, v \rangle \otimes_{alg} \mathbb{C} \langle S_v, v \leq T, v \rangle$ . But, for such a polynomial (in abstract variables), we have (in inserting  $1_{A_N}$  thanks to the case  $K = 2$  in (2.8) to use the definition of product in von Neumann algebra ultraproduct):

$$(P(W_t^N, \Upsilon_N) \# (W_t^N - W_Y^N))^\omega = P \# (S_t - S_Y)$$

and thus the stated orthogonality is obvious from the martingale property of matrix stochastic integrals of hermitian brownian motion.



Define  $G$  with  $f = G \circ (I_{t_1, \dots, t_k} * Id)$  and the associated  $X^{G, N, \Upsilon_N}$  from Theorem 5.4. We know the finite dimensional distribution of  $X^{G, N, \Upsilon_N}$  which is of the form assumed in proposition 2.7 so that, as before for  $S_s$ ,  $Y_s = (X_s^{G, N, \Upsilon_N})^\omega \in \mathcal{L}_P^\omega$  is actually  $Y_s \in \mathcal{M}_P^\omega$ .

**(i) First bounds on  $u_s = (b^{G, N, \Upsilon_N}(s, X^{G, N, \Upsilon_N}(s)))^\omega \in \mathcal{L}_P^\omega$ .**

We know that  $E(\|b^{G, N, \Upsilon_N}(s, X^{G, N, \Upsilon_N}(s))\|_2^2) \leq C$  independently of  $N, s$  so that  $\|u_s\|_2 \leq C$ .

Then from the bound in Theorem 5.4 we have for  $t, s \in [t_i, t_{i+1}]$ :

$$\|u_t - u_s\|_2^2 = \lim_{N \rightarrow \omega} E_P(\|b^{G, N, \Upsilon_N}(s, X^{G, N, \Upsilon_N}(s)) - b^{G, N, \Upsilon_N}(t, X^{G, N, \Upsilon_N}(t))\|_2^2) \leq C_4 \sqrt{|t - s|},$$

so that  $u$  is  $1/8$ -Hölder continuous on  $[t_i, t_{i+1}]$  as expected and especially  $u \in L_{ad}^\infty([0, 1], \mathcal{L}_P^\omega)$ .

**(ii) Limit of the value function along  $\omega$ .** As another consequence,  $\frac{1}{2} \int_0^1 \|u_s\|_2^2 ds$  can be approximated by Riemann sums, and the same approximations holds for  $\frac{1}{2} \int_0^1 \|b^{G, N, \Upsilon_N}(s, X^{G, N, \Upsilon_N}(s))\|_2^2 ds$  uniformly in  $N$  so that :

$$\lim_{N \rightarrow \omega} \frac{1}{2} \int_0^1 \|b^{G, N, \Upsilon_N}(s, X^{G, N, \Upsilon_N}(s))\|_2^2 ds = \frac{1}{2} \int_0^1 \|u_s\|_2^2 ds.$$

By convexity of  $G$  and Jensen's inequality, we also have :

$$E(G(\tau_{(X_{t_1}^{G, N, \Upsilon_N}, \dots, X_{t_k}^{G, N, \Upsilon_N})})) \geq G(E \circ \tau_{(X_{t_1}^{G, N, \Upsilon_N}, \dots, X_{t_k}^{G, N, \Upsilon_N})}).$$

Moreover, using again the uniform Hölder continuity to approximate integrals, we have  $Y_s = S_s + \int_0^s u_t dt$ , and the relation  $(u(X_{t_k}^{G, N, \Upsilon_N}) - 1)(X_{t_k}^{G, N, \Upsilon_N} - 4i) = 8i$  imply the corresponding relation in the ultraproduct so that  $u(Y_s) = (u(X_{t_k}^{G, N, \Upsilon_N}))^\omega$  and it is thus easy to see from the definition of product and trace in the ultraproduct that

$$\lim_{N \rightarrow \omega} d_2(E \circ \tau_{(X_{t_1}^{G, N, \Upsilon_N}, \dots, X_{t_k}^{G, N, \Upsilon_N}, \Upsilon_N)}, \tau_{(Y_{t_1}, \dots, Y_{t_k}, v)}) = 0.$$

Thus since  $G \in C^0(\mathcal{T}_{2,0}^c(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , one deduces :

$$\lim_{N \rightarrow \omega} E(G(\tau_{X^{G, N, \Upsilon_N}, \Upsilon_N})) \geq G(\tau_{(Y_{t_1}, \dots, Y_{t_k}, v)}).$$

Thus combining all our results and the formula from theorem 5.4, one gets :

$$\lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f(\tau_{W, \Upsilon_N})}) \geq f(\tau_{Y, v}) + \frac{1}{2} \int_0^1 \|u_s\|_2^2 ds.$$

To conclude with a bound bellow by  $\Lambda_{b, v}(f)$ , it only remains to check the two last conditions on  $u_s$ , namely  $u_s \in L^2(M)^m$  and  $W^*(v, Y_{t_1}, \dots, Y_{t_k})$  is a factor.

**(iii) Factoriality of  $W^*(v, Y_{t_1}, \dots, Y_{t_k})$ .**

For factoriality, by [D08, Th 4] and using the assumption  $m \geq 2$  or  $m = 1$  and  $B = W^*(v)$  diffuse, it suffices to check that  $(Y_{t_1}, \dots, Y_{t_k})$  have finite free Fisher information in the sense of [V5] relative to  $B = W^*(v)$ .

From the explicit knowledge of the law of  $(X_{t_1}^{G, N, \Upsilon_N}, \dots, X_{t_k}^{G, N, \Upsilon_N})$  in Theorem 5.4, one knows the classical score function of this random matrix model deduced from the density  $e^{-N^2 G(\tau_x, \Upsilon_N) - N^2 G_{2, t(f)}(\tau_x) - C}$  with respect to Lebesgue measure on Hermitian matrices. The score functions  $(-\Xi_{t_1}^{G, N}, \dots, -\Xi_{t_k}^{G, N})$  written in matrix form are thus (for  $i \geq 1$ , recall  $t_0 = 0$ ,  $G^N(x) = G(\tau_x, \Upsilon_N)$ ) given by:

$$\begin{aligned} \Xi_{t_i, l}^{G, N} &= \frac{N}{t_i - t_{i-1}} (X_{t_i, l}^{G, N, \Upsilon_N} - X_{t_{i-1}, l}^{G, N, \Upsilon_N}) + \frac{N}{t_{i+1} - t_i} (X_{t_i, l}^{G, N, \Upsilon_N} - X_{t_{i+1}, l}^{G, N, \Upsilon_N}) \\ &\quad + N^2 \mathcal{D}_i^l G^N(X_{t_1}^{G, N, \Upsilon_N}, \dots, X_{t_k}^{G, N, \Upsilon_N}). \end{aligned}$$

Consider a non-commutative polynomial  $P \in \mathbb{C}\langle X_1^1, \dots, X_1^m, X_2^1, \dots, X_k^m, u \rangle$  in  $km$  self-adjoint indeterminates and  $\mu\nu$  unitary indeterminates and let us write  $\partial_{X_j^l}$  the corresponding free difference quotient (with value 0 on  $u_j^i$ ).

As a consequence, one can write the integration by parts formula characterizing score functions:

$$(6.4) \quad E \left( \frac{1}{N^2} \text{Tr}(\Xi_{t_i,l}^{G,N} P(X_{t_1}^{G,N}, \dots, X_{t_k}^{G,N}, \Upsilon_N)) \right) = E \left( \frac{1}{N^2} (\text{Tr} \otimes \text{Tr})(\partial_{X_i^l} P(X_{t_1}^{G,N}, \dots, X_{t_k}^{G,N}, \Upsilon_N)) \right).$$

Note that  $N \mathcal{D}_i^l G^N = \frac{1}{\sqrt{N}} \mathcal{D}_i^l g_N$  and from the dimension independence of equation (3.3) obtained in lemma 5.3, one gets for some constants  $C, D > 0$  independent of  $N$ :

$$E \left( \left\| \frac{\Xi_{t_i,l}^{G,N}}{N} \right\|_2^2 \right) = E \left( \frac{1}{N^3} \text{Tr}((\Xi_{t_i,l}^{G,N})^* \Xi_{t_i,l}^{G,N}) \right) \leq C E \left( \sum_{i=1}^k \|X_{t_i}^{G,N}\|_2^2 \right) + D.$$

Thus  $\xi_{t_i}^l = \left( \frac{\Xi_{t_i,l}^{G,N}}{N} \right)^\omega \in \mathcal{L}_P^\omega$  is well-defined.

Using the second concentration result from proposition 2.7, one can take the limit  $N \rightarrow \omega$  in the right hand side of (6.4) and get the equation in  $\mathcal{L}_P^\omega$  (using also  $\xi_{t_i}^l = (\xi_{t_i}^l)^*$ ):

$$(6.5) \quad \langle \xi_{t_i}^l, P(Y_{t_1}, \dots, Y_{t_k}, v) \rangle = (\tau \otimes \tau) \left( (\partial_{X_i^l} P)(Y_{t_1}, \dots, Y_{t_k}, v) \right),$$

and this gives that  $E_{L^2(W^*(v, Y_{t_1}, \dots, Y_{t_k}))}(\xi_{t_i}^l)$  are exactly the conjugate variables in  $L^2(W^*(v, Y_{t_1}, \dots, Y_{t_k}))$  we were looking for.

**(iv) Limit along  $\omega$  of the value function  $h_t^{N, \Upsilon_n}(\sqrt{N} \cdot)/N^2$ .**

To prepare the identification of  $Y_t$  as strong solution of a free SDE that we will show to be a strong solution and therefore in  $L^2(M)$ , we first examine the limit along  $N \rightarrow \omega$  of the value function  $h_{t,N}(x, \Upsilon_N) := \frac{1}{N^2} h_t^{N, \Upsilon_N}(\sqrt{N}x)$ . We will be later able to get better convergence results, but, for now, we will be content of the  $\omega$  dependent result.

From lemma (5.3), the bounds for  $g_N$  in (3.3) are independent of  $d$  and thus from the proof of Proposition 3.5, so are the bounds for  $h_t^{N, \Upsilon_n}$  so that one gets constants  $C, D > 0$  such that for all  $N$ :

$$\begin{aligned} & |h_{t,N}(x_1, \dots, x_i, x_{i+1}, \Upsilon_N) - h_{t,N}(y_1, \dots, y_i, y_{i+1}, \Upsilon_N)| \\ & \leq \left( C \left( \sum_{K=1}^{i+1} \frac{1}{N} \text{Tr}(x_K^* x_K) \right)^{1/2} + C \left( \sum_{K=1}^{i+1} \frac{1}{N} \text{Tr}(y_K^* y_K) \right)^{1/2} + D \right) \left( \sum_{K=1}^{i+1} \frac{1}{N} \text{Tr}((x_K - y_K)^*(x_K - y_K)) \right)^{1/2} \end{aligned}$$

Thus applying this inequality to hermitian random variables  $X^N = (X_1^N, \dots, X_{i+1}^N)$ ,  $Y^N = (Y_1^N, \dots, Y_{i+1}^N)$ , taking expectation and using Cauchy-Schwartz inequality, we have:

$$\begin{aligned} & |E(h_{t,N}(X^N, \Upsilon_N)) - E(h_{t,N}(Y^N, \Upsilon_N))| \leq \left( \sum_{K=1}^{i+1} E \left( \frac{1}{N} \text{Tr}((X_K^N - Y_K^N)^2) \right) \right)^{1/2} \\ & \times \sqrt{E \left( 3C^3 \left( \sum_{K=1}^{i+1} \frac{1}{N} \text{Tr}((X_K^N)^2) \right) + 3C^2 \left( \sum_{K=1}^{i+1} \frac{1}{N} \text{Tr}((Y_K^N)^2) \right) + 3D^2 \right)} \end{aligned}$$

Thus considering  $X = (X^N)^\omega \in (\mathcal{L}_P^\omega)^{m(i+1)}$ , one can define :

$$h_t^\omega(X, v) = \lim_{N \rightarrow \omega} E(h_{t,N}(X^N, \Upsilon_N)).$$

Indeed, our previous inequality insures this is well-defined, namely, this does not depend on the way  $X = (X^N)^\omega = (Y^N)^\omega \in (\mathcal{L}_P^\omega)^{m(i+1)}$ . Actually, using the Lipschitzianity bound in variable  $\Upsilon_N$  in Theorem 5.4, one can show similarly that not only  $h_t^\omega(\cdot, v) : (\mathcal{L}_P^\omega)^{m(i+1)} \rightarrow \mathbb{R}$  is defined but also :

$$h_t^\omega(\cdot, \cdot) : (\mathcal{L}_P^\omega)^{m(i+1)} \times \mathcal{U}((M_N(\mathbb{C}))^\omega) \rightarrow \mathbb{R}.$$

Indeed, it suffices to note that a unitary in  $\mathcal{U}((M_N(\mathbb{C}))^\omega)$  can be represented by a sequence of unitaries (and  $\mathcal{U}((M_N(\mathbb{C}))^\omega)$  can even be identified with the ultraproduct of groups  $\mathcal{U}(M_N(\mathbb{C}))$ )

see e.g. [CL, Ex 2.11.6]) and the lipschitzness bound implies the function is well defined on the ultraproduct. We won't really use the second argument except at the fixed value  $v$ .

**(v) Regularity of the limit  $h_t^\omega$ .**

First,  $h_t^\omega$  is Lipschitz on bounded sets, from the inequality obtained by taking the limit of the one obtained for  $h_{t,N}$  in the previous point (iv):

$$|h_t^\omega(X, v) - h_t^\omega(Y, v)| \leq \|X - Y\|_2 \sqrt{3C^2\|X\|_2^2 + 3C^2\|Y\|_2^2 + 3D^2}.$$

Let us check that  $h_t^\omega(\cdot, v) : (\mathcal{L}_P^\omega)^{m(i+1)} \rightarrow \mathbb{R}$  for  $t_i < t \leq t_{i+1}$  is a convex function. Indeed, for  $\lambda \in [0, 1]$ , the relation

$$E(h_{t,N}(\lambda X^N + (1 - \lambda)Y^N), \Upsilon_N) \leq \lambda E(h_{t,N}(X^N), \Upsilon_N) + (1 - \lambda)E(h_{t,N}(Y^N), \Upsilon_N)$$

coming from the convexity of  $h_t$  goes to the limit  $N \rightarrow \omega$ .

From equation (3.2) and the uniformity of the constants in propositions 5.3 and 3.5, one gets:

$$h_t^\omega(X, v) \leq C(1 + \|X\|_2^2).$$

Finally, since equation (3.4) is checked with  $D_2 = 0$  and constants independent of  $N$ , from proposition 5.3, one deduces from proposition 3.5:

$$(6.6) \quad h_t^\omega(X + Y, v) - 2h_t^\omega(X, v) + h_t^\omega(X - Y, v) \leq C\|Y\|_2^2.$$

Thus from proposition 2.11,  $h_t^\omega$  is differentiable on  $(\mathcal{L}_P^\omega)^{m(i+1)}$  with Lipschitz derivative. Using (3.9), one also deduces (remembering that in our case  $D_{2,k,i} = 0$ ) that there are constants  $C, D$  such that for all  $t, t + s \in ]t_i, t_{i+1}[$ :

$$\begin{aligned} & |E(h_{t,N}(X_1, \dots, X_i, X, \Upsilon_N) - h_{t,N}(X_1, \dots, X_i, Y, \Upsilon_N) \\ & - h_{t+s,N}(X_1, \dots, X_i, X, \Upsilon_N) + h_{t+s,N}(X_1, \dots, X_i, Y, \Upsilon_N))| \\ & \leq \|Y - X\|_2 \sqrt{s} \times (C + D\|(X_1, \dots, X_i, Y)\|_2 + D\|(X_1, \dots, X_i, X)\|_2). \end{aligned}$$

Thus one deduces in taking the limit  $N \rightarrow \omega$ :

$$\begin{aligned} & |h_t^\omega(X_1, \dots, X_i, X, v) - h_t^\omega(X_1, \dots, X_i, Y, v) - h_{t+s}^\omega(X_1, \dots, X_i, X, v) + h_{t+s}^\omega(X_1, \dots, X_i, Y, v)| \\ & \leq \|Y - X\|_2 \sqrt{s} \times (C + D\|(X_1, \dots, X_i, Y)\|_2 + D\|(X_1, \dots, X_i, X)\|_2). \end{aligned}$$

and thus for all  $t, t + s \in ]t_i, t_{i+1}[$ ,  $s > 0$ :

$$(6.7) \quad \|\nabla_X h_t^\omega(X_1, \dots, X_i, X, v) - \nabla_X h_{t+s}^\omega(X_1, \dots, X_i, X, v)\|_2 \leq \sqrt{s}(C + 2D\|(X_1, \dots, X_i, X)\|_2).$$

**(vi)  $Y_t$  as solution of a free SDE.**

Note that we know by definition that

$$Y_s = S_s + \int_0^s u_t dt.$$

It remains to relate  $u_t$  to  $Y_t$ .

First recall that the function  $h_t^N$  from Theorem 5.4 is convex and thus we have the basic inequality for  $t_i < t \leq t_{i+1}$ ,  $H = (H_1, \dots, H_k)$  a vector of hermitian matrices:

$$(6.8) \quad \sum_{k=1}^m \text{Tr} \left( H_k [\mathcal{D}_{i+1}^k(h_t^{N, \Upsilon_N})](x_1, \dots, x_i, x) \right) \leq h_t^N(x_1, \dots, x_i, x + H) - h_t^{N, \Upsilon_N}(x_1, \dots, x_i, x).$$

thus applying this to the solution, one gets:

$$\begin{aligned} & \sum_{k=1}^m \frac{1}{N} \text{Tr} \left( H_k [b_k^{G, N, \Upsilon_N}(t, X_t^{G, N, \Upsilon_N}, \omega)] \right) \\ & \leq -\frac{1}{N^2} h_t^{N, \Upsilon_N}(\sqrt{N} X_{t_1}^{G, N, \Upsilon_N}(\omega), \dots, \sqrt{N} X_{t_i}^{G, N, \Upsilon_N}(\omega), \sqrt{N} X_t^{G, N, \Upsilon_N}(\omega)) \\ & \quad + \frac{1}{N^2} h_t^{N, \Upsilon_N}(\sqrt{N} X_{t_1}^{G, N, \Upsilon_N}(\omega), \dots, \sqrt{N} X_{t_i}^{G, N, \Upsilon_N}(\omega), \sqrt{N}(X_t^{G, N, \Upsilon_N}(\omega) + H)). \end{aligned}$$

Applying this to a random hermitian process  $H = H^N$ , taking expectations and the limit  $N \rightarrow \omega$ , one gets for  $H = (H^N)^\omega \in (\mathcal{L}_P^\omega)^m$ :

$$(6.9) \quad \sum_{k=1}^m \langle H_k, -u_{t,k} \rangle \leq h_t^\omega(Y_{t_1}, \dots, Y_{t_i}, Y_t + H, v) - h_t^\omega(Y_{t_1}, \dots, Y_{t_i}, Y_t, v)$$

By the characterization of (partial) subdifferentials, one gets  $-u_t \in \partial_{Y_t} h_t^\omega(Y_{t_1}, \dots, Y_{t_i}, Y_t, v)$ , and since  $h_t^\omega$  is differentiable, this subdifferential is single-valued equal to  $-u_t$ .

We now apply Theorem 4.3 to  $h_t^\omega$  restricted to  $L^2(M, \tau)^m \subset (\mathcal{L}_P^\omega)^m$  (filtration in which we noticed  $S_t$  is a free brownian motion. The regularity estimates obtained in step (v) enables to check most of the assumptions. Especially, (6.7) enables to check (4.5) with  $\alpha = 1/2$  and proposition 2.11 with (6.6) enables to check (4.6) with  $\beta = 1$ ,  $D = 0$  and the Gâteaux differentiability. Note that we also need to show what we will check in (ix) that, with the notation of Theorem 4.3,  $\nabla_{i+1} h_t^\omega(X) \in L^2(M_t)$  if  $X \in L^2(M_t)$ . Consider  $P_s$  the projection on  $L^2(M_s, \tau)^m$  and  $P$  the projection on  $L^2(M, \tau)^m$ .

Projecting the equation for  $Y_t$  and using continuity of  $P$  to make it commute with the integral, one gets:

$$P(Y_s) = S_s - \int_0^s P(\nabla_{Y_t} h_t^\omega(Y_{t_1}, \dots, Y_{t_i}, Y_t, v)) dt.$$

From the orthogonality property (6.3), one deduces  $P(Y_s) = P_s(Y_s)$  and

$$P(\nabla_{Y_t} h_t^\omega(Y_{t_1}, \dots, Y_{t_i}, Y_t, v)) = P_t(\nabla_{Y_t} h_t^\omega(Y_{t_1}, \dots, Y_{t_i}, Y_t, v)).$$

From the existence of the solution  $Z_t$  in  $L^2(M, \tau)^m$  of the SDE in Theorem 4.3 associated to  $h_t^\omega$  restricted to  $L^2(M, \tau)^m$  and we know that  $Z_t$  satisfies :

$$Z_s = S_s - \int_0^s P_t(\nabla_{Z_t} h_t^\omega(Z_{t_1}, \dots, Z_{t_i}, Z_t, v)) dt.$$

since the gradient of the restriction of  $h_t^\omega$  is the same as the projection by  $P_t$  of the gradient of  $h_t^\omega$ .

It suffices to argue as in the proof of the uniqueness in Theorem 4.3 in order to get  $Z_s = P_s(Y_s)$ . But we actually want to show  $Z_s = Y_s$  in checking we can remove  $P_t$  from the equation for  $Z$ . This will require an alternative formula for  $\nabla h_t^\omega$ .

**(vii) Formula for  $\nabla h_t^\omega$ .**

First note that by the fundamental theorem of calculus, the identity  $\mathcal{D}_{i+1}^k(h_{t,N})(x_1, \dots, x_i, x, \Upsilon_N) = \frac{\sqrt{N}}{N^2} \mathcal{D}_{i+1}^k(h_t^{N, \Upsilon_N})(\sqrt{N}x_1, \dots, \sqrt{N}x_i, \sqrt{N}x)$  and the bounds on gradients of  $h_{t,N}$  from (v) and proposition 2.11, one gets:

$$\begin{aligned} & |h_{t,N}(x_1, \dots, x_i, x, \Upsilon_N) - h_{t,N}(x_1, \dots, x_i, y, \Upsilon_N) \\ & - \sum_{k=1}^m \frac{1}{N\sqrt{N}} \text{Tr}(\mathcal{D}_{i+1}^k(h_t^{N, \Upsilon_N})(\sqrt{N}x_1, \dots, \sqrt{N}x_i, \sqrt{N}x)(y - x)_k) \Big| \leq \frac{C}{2} \|y - x\|_2^2. \end{aligned}$$

with the constant  $C$  of equation (6.6).

We now want to use, for  $t_i < t \leq t_{i+1}$  fixed, the last formula in corollary 5.6. We thus consider a random variable  $X^N = (X_1^N, \dots, X_i^N, X_{i+1}^N)$  a vector of hermitian matrices and  $\mathcal{F}_t$ -measurable (i.e. adapted to the filtration of hermitian brownian motion), and consider the solution  $X^{G,N,t}(s)$  in this corollary starting from  $X^{G,N, \Upsilon_N,t}(s) = X_{i+1}^N$  which is independent of the noise appearing in the equation. As before in (i), one defines  $Y_s^t = (X^{G,N, \Upsilon_N,t}(s))^\omega \in \mathcal{L}_P^\omega$  and then as in (i)  $u_s^t = (b^{G,N, \Upsilon_N,t}(u, X^{G,N, \Upsilon_N,t}(u)))^\omega \in \mathcal{L}_P^\omega$ . From the same a priori Hölder continuity bounds, we know for  $s \geq t$ :

$$Y_s^t = (X_{i+1}^N)^\omega + \int_t^s dv u_v^t + (S_s - S_t).$$

Using the convexity of  $h^\omega$ , one deduces as in (vi) that for  $t_I < s < t_{I+1}$  :

$$u_s^t = \nabla_{Y_s^t} h_s^\omega(X_1, \dots, X_i, Y_{t_i}^t, \dots, Y_{t_I}^t, Y_s^t, v)$$

Similarly, one takes

$$v_l^t = \sum_{j=i+1}^k \left( \frac{1}{\sqrt{N}} (\mathcal{D}_j^l g_N(\sqrt{N}X_1^N, \dots, \sqrt{N}X_i^N, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_{i+1}, X_{i+1}^N), \dots, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_k, X_{i+1}^N))) \right)^\omega \in \mathcal{L}_P^\omega$$

(the a priori boundedness from (3.3)). First note that as for  $h_{N,t}$ ,

$$G^\omega(X^\omega, v) = \lim_{N \rightarrow \omega} E\left(\frac{1}{N^2} g_N(\sqrt{N}X^N)\right)$$

defines a well-defined convex function on the ultraproduct (this is actually the special case  $G^\omega = h_1^\omega$ ). Considering the convexity relation:

$$\begin{aligned} & \sum_{l=1}^m \frac{1}{N} \text{Tr} \left( H_l \left[ \sum_{j=i+1}^k \left( \frac{1}{\sqrt{N}} (\mathcal{D}_j^l g_N(\sqrt{N}X_1^N, \dots, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_{i+1}, X_{i+1}^N), \dots, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_k, X_{i+1}^N))) \right) \right] \right) \\ & \leq -\frac{1}{N^2} g_N(\sqrt{N}X_1^N, \dots, \sqrt{N}X_i^N, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_{i+1}, X_{i+1}^N), \dots, \sqrt{N}X^{G,N,\Upsilon_N,t}(t_k, X_{i+1}^N)) \\ & \quad + \frac{1}{N^2} g_N(\sqrt{N}X_1^N, \dots, \sqrt{N}[X^{G,N,\Upsilon_N,t}(t_{i+1}, X_{i+1}^N) + H], \dots, [\sqrt{N}X^{G,N,\Upsilon_N,t}(t_k, X_{i+1}^N) + H]). \end{aligned}$$

From which, replacing  $H$  by  $H^N$  random and then defining  $H = (H^N)^\omega$ , taking the expectation and the limit  $N \rightarrow \omega$  one deduces:

$$\sum_{l=1}^m \langle H_l, v_l^t \rangle \leq -G^\omega(X_1^\omega, \dots, X_i^\omega, Y_{t_{i+1}}^t, \dots, Y_{t_k}^t, v) + G^\omega(X_1^\omega, \dots, X_i^\omega, Y_{t_{i+1}}^t + H, \dots, Y_{t_k}^t + H, v).$$

and as a consequence one deduces

$$v^t = \sum_{j=i+1}^k \nabla_{Y_{t_j}^t} G^\omega(X_1^\omega, \dots, X_i^\omega, Y_{t_{i+1}}^t, \dots, Y_{t_k}^t, v).$$

Now, one can take the limit  $N \rightarrow \omega$  in the first relation obtained in this point (vii) for  $X_l^\omega, Y, Y_l$   $\mathcal{L}_{P,t}^\omega$ -measurable after taking expectations and using the last equation in corollary 5.6 and get:

$$\begin{aligned} & \left| h_t^\omega(X_1^\omega, \dots, X_i^\omega, X_{i+1}^\omega, v) - h_t^\omega(Y_1, \dots, Y_i, Y, v) - \sum_{l=1}^m \langle E_{\mathcal{L}_{P,t}^\omega}(v_l^t), (Y - X_{i+1}^\omega)_l \rangle \right| \\ & \leq \frac{C}{2} \|X_{i+1}^\omega - Y\|_2^2 \end{aligned}$$

and thus

$$\nabla_{X_{i+1}^\omega} h_t^\omega(X_1^\omega, \dots, X_i^\omega, X_{i+1}^\omega, v) = E_{\mathcal{L}_{P,t}^\omega} \left( \sum_{j=i+1}^k \nabla_{Y_{t_j}^t} G^\omega(X_1^\omega, \dots, X_i^\omega, Y_{t_{i+1}}^t, \dots, Y_{t_k}^t, v) \right).$$

**(viii) Computation of  $G^\omega, \nabla_X G^\omega$ .** Finally, we will need a way to compute  $G^\omega$ , since they appear in the above formula for  $\nabla_{X_{i+1}^\omega} h_t^\omega$ . We start with the value on  $(X_1, \dots, X_k) \in L^2(M)$ ,  $X_i = X_i^*$  such that there is  $X_0 = X_0^* \in M$  with  $W^*(v, X_0, X_1, \dots, X_k) = W^*(v, X_0, u(X_1), \dots, u(X_k))$  is a factor. Then for any model with  $(X_i^N)^\omega = X_i$  which implies as above  $(u(X_i^N))^\omega = u(X_i)$ . Let  $Y = (X_0, u(X_1) + u(X_1)^*, i(u(X_1) - u(X_1)^*), \dots, i(u(X_k) - u(X_k)^*)), Y^N$  similarly,  $\tau_{Y,v} \in \mathcal{S}_R^{(2k+1)m} \star \mathcal{T}(\mathcal{F}_\mu^\nu)$  is extremal and we can apply proposition 2.10 so that from  $\lim_{N \rightarrow \omega} d_{2,0}(E \circ \tau_{Y^N, \Upsilon_N}, \tau_{Y,v}) = 0$ , one deduces  $\lim_{N \rightarrow \omega} E(d_{2,0}(\tau_{Y^N, \Upsilon_N}, \tau_{Y,v})) = 0$  and then in rewriting variables in terms of  $X_1, \dots, X_k$  and using the lipschitzness of  $G$ :

$$G^\omega(X_1, \dots, X_k, v) = G(\tau_{X_1, \dots, X_k, v}) \text{ if } W^*(X_0, X_1, \dots, X_k, v) \text{ factor.}$$

We now establish a variant for the derivative. From the approximation property in the definition of  $\mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , for  $\epsilon > 0$ , one can fix  $P_1^1, \dots, P_L^m \in \mathcal{F}_k^m * \mathcal{F}_\mu^\nu$ ,  $f_1^1, \dots, f_L^m, g_{1,1,1}, \dots, g_{k,m,m} \in$

$C^0(\mathcal{T}_{2,0}(\mathcal{F}_k^m * \mathcal{F}_\mu^\nu), d_{2,0})$ ) such that inserting in the first equation of (vii) for  $t = 1$ ,  $X = (x_1, \dots, x_k)$

$$\left| G(\tau_{x_1, \dots, x_k, \Upsilon_N}) - G(\tau_{x_1, \dots, x_o+y, \dots, x_k, \Upsilon_N}) - \sum_{K=1}^m \frac{1}{N} \text{Tr} \left( \left[ \sum_{i=1}^L P_i^K(u(X), \Upsilon_N) f_i^K(\tau_X, \Upsilon_N) + \sum_{i=1}^m \sum_{j=1}^k x_j^{(i)} g_{j,i,K}(\tau_X, \Upsilon_N) \right] y_K \right) \right| \leq \frac{C}{2} \|y\|_2^2 + \epsilon.$$

Taking the limit  $N \rightarrow \omega$ , treating each function of trace by concentration as before, and finally letting  $\epsilon \rightarrow 0$ , one gets if  $W^*(X_0, X_1, \dots, X_k, v)$  factor, then for all  $o$ :  $\nabla_{X_o} G^\omega(X_1, \dots, X_k) = \nabla_{X_o} G(\tau_{X_1, \dots, X_k}) \in L^2(W^*(v, X_1, \dots, X_k))^m$ .

Finally, let us extend this to any  $(X_1, \dots, X_k) \in L^2(M)$ ,  $X_i = X_i^*$ .

By Clarck Ocone's formula [BS], for  $B = W^*(v)$ , each  $X_i = E_B(X_i) + \int_0^1 U_s dS_s$  thus there is  $X_{i,\epsilon} = E_B(X_i) + \int_{\eta(\epsilon)}^1 U_s dS_s$  with  $\|X_{i,\epsilon} - X_i\| \leq \epsilon$ . Moreover if  $X_{0,\epsilon} = S_{\eta(\epsilon)}$ , let us check that  $W^*(X_{0,\epsilon}, X_{1,\epsilon}, \dots, X_{k,\epsilon}, v)$  is a factor and thus satisfies our previous assumption. Indeed from [V5],  $X_{0,\epsilon}$  have bounded first and second order conjugate variables, thus, if  $m \geq 2$  (the case  $m = 1$  is similar in taking two increments  $S_{\eta(\epsilon)/2}, S_{\eta(\epsilon)} - S_{\eta(\epsilon)/2}$  instead of one) from [D08, Rmk 11],  $X_{0,\epsilon}$  is a non- $\Gamma$  set thus a non-amenability set ([Co76], see e.g. [DI, Def 2.4, lemma 2.10]) and  $W^*(X_{0,\epsilon}, X_{1,\epsilon}, \dots, X_{k,\epsilon}, v) \subset W^*(X_{0,\epsilon}) * A = M$  is contained in some free product (using the free brownian motion property and  $v$  free from  $X_{0,\epsilon}$ ). For any  $Z$  in the center,  $E_{W^*(X_{0,\epsilon})}(Z)$  is in the center of  $W^*(X_{0,\epsilon})$  which is a factor (since it is non- $\Gamma$ , even a free group factor in our example) thus  $Z - \tau(Z) = Z - E_{W^*(X_{0,\epsilon})}(Z) \in L^2(W^*(X_{0,\epsilon}) * A) \ominus L^2(W^*(X_{0,\epsilon}))$  which is well-known to be a (countable) direct sum of coarse correspondences over  $W^*(X_{0,\epsilon})$ . But the non-amenability set property implies :

$$\|Z - E_{W^*(X_{0,\epsilon})}(Z)\| \leq K \sum_{i=1}^m \|[X_{0,\epsilon}, Z - E_{W^*(X_{0,\epsilon})}(Z)]\| = 0.$$

Thus  $Z = \tau(Z)$  and we deduce the expected factoriality. Thus, from the first case, we deduce:  $G^\omega(X_{1,\epsilon}, \dots, X_{k,\epsilon}, v) = G(\tau_{X_{1,\epsilon}, \dots, X_{k,\epsilon}, v})$  and  $\nabla_{X_{o,\epsilon}} G^\omega(X_{1,\epsilon}, \dots, X_{k,\epsilon}, v) = \nabla_{X_{o,\epsilon}} G(\tau_{X_{1,\epsilon}, \dots, X_{k,\epsilon}, v})$ .

Since  $G, G^\omega$  and their gradients are continuous (even Lipschitz) by assumption, one deduces in taking  $\epsilon \rightarrow 0$  that for any  $(X_1, \dots, X_k, X) \in L^2(M)^{k+1}$

$$(6.10) \quad G^\omega(X_1, \dots, X_k, v) = G(\tau_{X_1, \dots, X_k, v}),$$

$$(6.11) \quad \nabla_{X_o} G^\omega(X_1, \dots, X_k, v) = \nabla_{X_o} G(\tau_{X_1, \dots, X_k, v}) \in L^2(W^*(X_1, \dots, X_k, v))^m.$$

**(ix) Induction and Picard iterations for  $Y_s^t, \nabla_X h_t^\omega$ .** We are now ready to use the formula obtained in (vii) to extend the result of (viii) (beyond the case  $t = 1$ ) and obtain for any  $(X_1, \dots, X_k, X) \in L^2(M_t)^{k+1}$ ,  $t_i < t \leq t_{i+1}$

$$(6.12) \quad \nabla_X h_t^\omega(X_1, \dots, X_i, X, v) \in L^2(W^*(X_1, \dots, X_i, X, v))^m,$$

$$(6.13) \quad Y_s^t(X_1, \dots, X_i, X, v) \in L^2(W^*(X_1, \dots, X_i, X, S_u - S_t, u \in [t, s], v))^m, \quad s \in [t, 1].$$

We now have a closed system of equations determining  $Y_s^t$  in terms of  $\nabla_X h_t^\omega$  and vice-versa.

We will analyse it in an induction on  $k - i$  starting at  $k - i = 0$ ,  $t_k < t \leq 1$ . This is the trivial case since for such  $t$ ,  $\nabla_X h_t^\omega = 0$  (the sum is empty in the last formula from (vii) and thus  $Y_s^t(X_1, \dots, X_k, X) = X + S_s - S_t$ . Similarly later, we can disregard the case  $s > t_k$ , since  $Y_s^t(X_1, \dots, X_k, X, v) = Y_{t_k}^t(X_1, \dots, X_k, X, v) + S_s - S_{t_k}$ .

The real initialization case is  $k - i = 1$ ,  $t_{k-1} \leq t \leq t_k$ . In this case, the equations in (vii) can be written:

$$(6.14) \quad \begin{aligned} \nabla_X h_t^\omega(X_1, \dots, X_{k-1}, X, v) &= E_{\mathcal{L}_{P,t}^\omega} \left[ \nabla_{Y_{t_k}^t} G^\omega(X_1, \dots, X_{k-1}, Y_{t_k}^t(X_1, \dots, X_{k-1}, X, v)) \right], \\ Y_s^t(X_1, \dots, X_{k-1}, X, v) &= X + \int_t^s dv \nabla_{Y_v^t} h_v^\omega(X_1, \dots, X_{k-1}, Y_v^t(X_1, \dots, X_{k-1}, X, v)) + (S_s - S_t). \end{aligned}$$



Note that we know from (6.6) that  $\nabla_X h_t^\omega$  is  $C$ -Lipschitz and thus from the definition and an application of Gronwall's lemma,  $Y_s^t$  is Lipschitz with constant  $(1 + C)e^{C(s-t)} \leq (1 + C)e^C$ .

We will first consider the case  $t \in [t_k - \epsilon, t_k]$  for  $\epsilon > 0$  to be chosen small enough later. On this interval, we will approximate our system of equations by a Picard iteration. We define inductively for  $X, X_m \in (\mathcal{L}_{P,t}^\omega)^m$ ,  $s \in [t, t_k]$ ,  $t_k - \epsilon < t \leq t_k$

$$Y_s^{t,(0)}(X) = Y_s^{t,(0)}(X_1, \dots, X_{k-1}, X, v) = X + (S_s - S_t),$$

for  $l \geq 0$  :

$$DH_t^{(l)}(X_1, \dots, X_{k-1}, X, v) = E_{\mathcal{L}_{P,t}^\omega} \left( \sum_{j=i+1}^k \nabla_{Y_{t_j}^{t,(l)}} G^\omega(X_1, \dots, X_{k-1}, Y_{t_k}^{t,(l)}(X_1, \dots, X_{k-1}, X, v), v) \right)$$

and for  $l \geq 1$ :

$$\begin{aligned} Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v) &= X + (S_s - S_t) \\ &+ \int_t^s dv DH_v^{(l-1)}(X_1, \dots, X_{k-1}, Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v), v). \end{aligned}$$

We first want to check by induction on  $l$  that the integral above is well-defined and

$$Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v) \in L^2(W^*(v, X_1, \dots, X_i, X, S_u - S_t, u \in [t, s]))^m$$

if  $X, X_m \in (L^2(M_t))^m$ . The initialization  $l = 0$  is obvious and then if the result is true at level  $l$ ,  $X_1, \dots, X_{k-1}, Y_{t_k}^{t,(l)}(X_1, \dots, X_{k-1}, X, v) \in L^2(M_{t_k})$  and applying (6.11), so is the application  $\nabla_{Y_{t_j}^{t,(l)}} G^\omega$  in the definition of  $DH_t^{(l)}(X_1, \dots, X_{k-1}, X, v)$ . But on this space  $L^2(M_{t_k})$ , from the orthogonality relation (6.3), one deduces  $E_{\mathcal{L}_{P,t}^\omega} = E_{L^2(M_t)}$  and thus from freeness we even deduce  $DH_t^{(l)}(X_1, \dots, X_{k-1}, X) \in L^2(W^*(X_1, \dots, X_i, X, v))$ . This implies the right continuity in  $v$  of the integrand proved inductively simultaneously in using the lipschitzness of  $\nabla G^\omega$  and the right continuity of the filtration  $L^2(M_t)$  (coming for instance from the free Clark-Ocone's formula). especially, at each inductive step, the integral is well-defined. Applying this with the induction hypothesis to the integrand defining  $Y_s^{t,(l+1)}$ , one obtains the induction step.

Let us compute bounds on Lipschitzness constants :  $\|Y_s^{t,(0)}\|_{Lip} \leq 1$  and knowing that  $\nabla G^\omega$  is  $C$ -Lipschitz, we have by composition for  $t_k - \epsilon \leq t \leq t_k$

$$\begin{aligned} &\|DH_t^{(l)}(X_1, \dots, X_{k-1}, X, v) - DH_t^{(l)}(Z_1, \dots, Z_{k-1}, Z), v\|_2 \\ &\leq C(1 + \|Y_t^{t,(l)}\|_{Lip}^2)^{1/2} \|(X_1, \dots, X_{k-1}, X) - (Z_1, \dots, Z_{k-1}, Z)\|_2 \end{aligned}$$

And similarly

$$\|Y_s^{t,(l)}\|_{Lip} \leq 1 + \int_t^s dv \|DH_v^{(l-1)}\|_{Lip} (1 + \|Y_v^{t,(l-1)}\|_{Lip}^2)^{1/2}.$$

If  $(1 + (2 + C)^2 e^{2C})C\epsilon < 1$ , an immediate induction yields for  $s, t$  as above in the definition  $\|Y_s^{t,(l)}\|_{Lip} \leq 1 + (1 + (2 + C)^2 e^{2C})C\epsilon \leq 2 \leq E := (2 + C)e^C$ , (where the supplementary constant with  $E - 2$  is to keep the choice of  $\epsilon$  independent of our next induction step) and  $\|DH_s^{(l-1)}\|_{Lip} \leq D := \sqrt{1 + (2 + C)^2 e^{2C}}C$ . Thus, one can bound the increments:

$$\begin{aligned} &\|DH_t^{(l)}(X_1, \dots, X_{k-1}, X, v) - DH_t^{(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\ &\leq C\|Y_{t_k}^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_{t_k}^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2, \end{aligned}$$

Thus similarly and using the relation  $Y_{t_k}^{v,(l)}(X_1, \dots, X_{k-1}, Y_v^{t,(l)}(X_1, \dots, X_{k-1}, X, v), v) = Y_{t_k}^{t,(l)}(X_1, \dots, X_{k-1}, X, v)$  obtained by induction, one gets:

$$\begin{aligned}
& \|Y_s^{t,(l+1)}(X_1, \dots, X_{k-1}, X, v) - Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& \leq \int_t^s dv \|DH_v^{(l)}(X_1, \dots, X_{k-1}, Y_v^{t,(l)}(X_1, \dots, X_{k-1}, X, v), v) \\
& \quad - DH_v^{(l-1)}(X_1, \dots, X_{k-1}, Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v))\|_2 \\
& \leq C \int_t^s dv \|Y_{t_k}^{v,(l)}(X_1, \dots, X_{k-1}, Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v), v) \\
& \quad - Y_{t_k}^{v,(l-1)}(X_1, \dots, X_{k-1}, Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v), v)\|_2 \\
& + \int_t^s dv D \|Y_v^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& \leq \int_t^s dv C \|Y_{t_k}^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_{t_k}^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& + \int_t^s dv (EC + D) \|Y_v^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2.
\end{aligned}$$

Thus taking a supremum, one gets

$$\begin{aligned}
& \sup_{t < s \in [t_k - \epsilon, t_k]^2} \|Y_s^{t,(l+1)}(X_1, \dots, X_{k-1}, X, v) - Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& \leq (EC + C + D)\epsilon \sup_{t < v \in [t_k - \epsilon, t_k]^2} \|Y_v^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& \leq [(EC + C + D)C\epsilon]^l \sup_{t < v \in [t_k - \epsilon, t_k]^2} \|Y_v^{t,(1)}(X_1, \dots, X_{k-1}, X, v) - Y_v^{t,(0)}(X_1, \dots, X_{k-1}, X, v)\|_2.
\end{aligned}$$

If we moreover assume  $e^C(EC + C + D)C\epsilon < 1$  (again the supplementary  $e^C$  for the next induction step) the series above converges and thus  $Y_s^{t,(l+1)}(X_1, \dots, X_{k-1}, X, v) \rightarrow Y_s^{t,(\infty)}(X_1, \dots, X_{k-1}, X, v)$  uniformly on  $s < t \in [t_k - \epsilon, t_k]^2$ , and as a consequence  $DH_t^{(l)}(X_1, \dots, X_{k-1}, X, v) \rightarrow DH_t^{(\infty)}(X_1, \dots, X_{k-1}, X, v)$  uniformly on  $t \in [t_k - \epsilon, t_k]$ .

To identify the limits, one uses bounds similar to those used for increments:

$$\begin{aligned}
& \|DH_t^{(l)}(X_1, \dots, X_{k-1}, X, v) - \nabla_X h_t^\omega(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& \leq C \|Y_{t_k}^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_{t_k}^t(X_1, \dots, X_{k-1}, X, v)\|_2,
\end{aligned}$$

and using this time the semigroup property coming from uniqueness of the solution of the SDE  $Y_{t_k}^v(X_1, \dots, X_{k-1}, Y_v^t(X_1, \dots, X_{k-1}, X, v), v) = Y_{t_k}^t(X_1, \dots, X_{k-1}, X, v)$  (note that since  $S_t$  is not a free brownian motion in the considered filtration of the ultraproduct, this is not an application of Theorem 4.3, but the uniqueness part would work for any driving process not necessarily free brownian motion), one gets:

$$\begin{aligned}
& \sup_{t < s \in [t_k - \epsilon, t_k]^2} \|Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_s^t(X_1, \dots, X_{k-1}, X, v)\|_2 \\
& \leq [(EC + C + D)C\epsilon]^l \sup_{t < v \in [t_k - \epsilon, t_k]^2} \|Y_v^{t,(1)}(X_1, \dots, X_{k-1}, X, v) - Y_v^t(X_1, \dots, X_{k-1}, X, v)\|_2 \xrightarrow{l \rightarrow \infty} 0.
\end{aligned}$$

From the limit, we have thus deduced (6.12) and (6.13) for  $t < s \in [t_k - \epsilon, t_k]^2$ .

Now with the same fixed  $\epsilon$  we go on in checking the relation for  $t \in [t_k - 2\epsilon, t_k - \epsilon]^2, s \in ]t, t_k]$ . We define  $DH_t^{(l)}$  with the same formula but we now consider: for  $l \geq 1$ :

$$\begin{aligned} Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v) &= X + (S_s - S_t) \\ &+ \int_t^{s \wedge (t_k - \epsilon)} dv DH_v^{(l-1)}(X_1, \dots, X_{k-1}, Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v), v) \\ &+ \int_{s \wedge (t_k - \epsilon)}^s dv \nabla_X h_v^\omega(X_1, \dots, X_{k-1}, Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v), v). \end{aligned}$$

and similarly for  $l = 0$  without integrals. Since we have all the properties we need from the previous step for the last integral, we can use the a priori Lipschitz bounds on those terms to obtain the same bounds as before since the supplementary  $E - 2$  was introduced with this goal. Note also that for  $S \in [t_k - \epsilon, t_k]$ :

$$\begin{aligned} &\sup_{s \in [t, S]} \|Y_s^{t,(l+1)}(X_1, \dots, X_{k-1}, X, v) - Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\ &\leq [(EC + C + D)C\epsilon] \sup_{v \in [t, t_k]^2} \|Y_v^{t,(l)}(X_1, \dots, X_{k-1}, X, v) - Y_v^{t,(l-1)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\ &+ C \int_{t_k - \epsilon}^S du \sup_{s \in [t, u]} \|Y_s^{t,(l+1)}(X_1, \dots, X_{k-1}, X, v) - Y_s^{t,(l)}(X_1, \dots, X_{k-1}, X, v)\|_2 \\ &\leq [e^C (EC + C + D)C\epsilon]^l \sup_{v \in [t, t_k]} \|Y_v^{t,(1)}(X_1, \dots, X_{k-1}, X, v) - Y_v^{t,(0)}(X_1, \dots, X_{k-1}, X, v)\|_2. \end{aligned}$$

The key inequality comes from Gronwall's lemma, the first inequality uses the same argument as in the initialization. The same choice of  $\epsilon$  thus enables to get convergence when  $l \rightarrow \infty$  with the expected terms. An obvious induction then concludes to (6.12) and (6.13) for  $t < s \in [t_{k-1}, t_k]^2$ .

Assume (6.12) and (6.13) at induction step  $k - i \geq 1$  thus for  $t \in [t_i, t_{i+1}], s \in ]t, t_k]$ . Let us prove the case  $t \in [t_{i-1}, t_i], s \in ]t, t_k]$ .

We start by the case  $s \in ]t, t_i]$  that will be similar to our previous initialization. Indeed, in this case, the equations from (vii) can be written:

$$\begin{aligned} &\nabla_X h_t^\omega(X_1, \dots, X_{k-1}, X, v) \\ &= E_{\mathcal{L}_{P,t}^\omega} \left[ \nabla_{Y_{t_k}^t} G^\omega(X_1, \dots, X_{i-1}, Y_{t_i}^t(X_1, \dots, X_{i-1}, X, v), \dots, Y_{t_k}^t(X_1, \dots, X_{i-1}, X, v), v) \right] \\ &= E_{\mathcal{L}_{P,t}^\omega} \left[ \nabla_{Y_{t_k}^t} G^\omega(X_1, \dots, X_{i-1}, Y_{t_i}^t(X_1, \dots, X_{i-1}, X, v), \right. \\ &\quad \left. \dots, Y_{t_k}^{t_i}(X_1, \dots, X_{i-1}, Y_{t_i}^t(X_1, \dots, X_{i-1}, X, v), v), v) \right], \end{aligned}$$

and the same equation (6.14) as before for  $Y_s^t$ . Since we already treated the case of  $Y_{t_k}^{t_i}$  we know for this map the same properties as for  $\nabla_{Y_{t_k}^t} G^\omega$  and it only suffices to take  $\epsilon > 0$  smaller so that the same Picard iteration as for the initialization case concludes. The details are left to the reader. The really new case in this induction step is for  $s \in ]t_i, t_k]$ . In this case the equations from (vii) for  $\nabla_X h_t^\omega$  is the same and  $Y_s^t(X_1, \dots, X_{i-1}, X, v) = Y_{t_i}^{t_i}(X_1, \dots, X_{i-1}, Y_{t_i}^t(X_1, \dots, X_{i-1}, X, v), v)$  and the composition of previous cases concludes.

We thus deduced (6.12) and (6.13) in all the expected cases.

An examination of the proof even shows that  $\nabla_X h_t^\omega$  does not depend on  $\omega$  on  $L^2(M_t)$  as a consequence of the construction above and the independence of  $\omega$  of  $\nabla_X G^\omega$  (case  $t = 1$  in (viii)). We won't use this remark.

### (x) Conclusion

One can now turn back to the solution  $Z_s \in L^2(M_s)^m$  built in step (vi), point (ix) applies so that  $\nabla_{Z_t} h_t^\omega(Z_{t_1}, \dots, Z_{t_i}, Z_t, v) \in L^2(W^*(Z_{t_1}, \dots, Z_{t_i}, Z_t, v))^m \subset L^2(M_t)$  and thus

$P_t(\nabla_{Z_t} h_t^\omega(Z_{t_1}, \dots, Z_{t_i}, Z_t, v) = \nabla_{Z_t} h_t^\omega(Z_{t_1}, \dots, Z_{t_i}, Z_t, v)$  and  $Z_t$  satisfies the SDE:

$$Z_s = S_s - \int_0^s \nabla_{Z_t} h_t^\omega(Z_{t_1}, \dots, Z_{t_i}, Z_t, v) dt.$$

This is the same SDE as for  $Y_s$ , and thus by the uniqueness exactly as in the proof of Theorem (4.3), one deduces  $Y_s = Z_s \in L^2(M_s)^m$ . Thus by differentiation in time  $u_t \in (L^2(M_t, \tau))^m$  (for almost every time, this uses the right continuity of the filtration of free brownian motion) and thus  $u \in L_{ad}^\infty([0, 1], (L^2(M, \tau))^m)$ . Thus, we can consider  $Y$  as one of the processes entering in the infimum for  $\Lambda_{b,v}(f)$  and we deduce :

$$f(\tau_Y) + \frac{1}{2} \int_0^1 \|u_s\|^2 ds \geq \Lambda_{b,v}(f),$$

and, from (ii), (6.2) is satisfied.

**Step 3 :** Upper bound for  $f \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \cup \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}), p \in [2, \infty[$ .

Consider an ultrafilter  $\omega \in \beta\mathbb{N} - \mathbb{N}$ . It suffices to show that :

$$\lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \tau_N)}) \leq \Lambda_{b,v}(f).$$

Thus take  $u \in \mathcal{P}_{ad,1/8loc,t(f),pl} \cap \mathcal{P}_{ad,factor,t(f)}$  so that, from lemma 6.1, it suffices to show:

$$(6.15) \quad \lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \tau_N)}) \leq \frac{1}{2} \int_0^1 \|u_s\|^2 ds + f(S + \int_0^\cdot u_s ds).$$

We can see  $u_s \in L^2(M)^m \subset (\mathcal{L}_P^\omega)^m$  and we have a set of times

$$\mathbf{s}(u) = \{s_1^1 < \dots < s_n^1 < \dots s_\omega^1 = t_1 < s_1^2 < \dots < s_\omega^2 = t_2 < \dots s_\omega^k = t_k < s_1^{k+1} < \dots s_l^{k+1}\}$$

enabling to realize the piecewise linear property.

We want to define a continuous adapted process  $V_s^N \in M_N(L^2(\mathbb{W}_{sa,N}, \gamma_N))$ .

Take  $U_{s_k^l}^N \in M_N(L^2(W_t^N, t \leq s_{k-1}^l))$  such that

$$(U_{s_k^l}^N)^\omega = \int_0^{s_k^l} u_s ds.$$

One can assume  $\|U_{s_k^l}^N\|_2 \leq \|\int_0^{s_k^l} v_s ds\|_2$ .

For each  $l$  for  $k$  large enough, one can take  $U_{s_k^l}^N \in M_N(L^\infty(W_t^N, t \leq s_{k-1}^l))^m$  since then  $\int_0^{s_k^l} v_s ds \in (\mathcal{M}_P^\omega)^m$  and then assume instead the operator norm bound  $\|U_{s_k^l}^N\| \leq \|\int_0^{s_k^l} v_s ds\| + 1 \leq C + 1$  where  $C$  is given by the definition of  $\mathcal{P}_{ad,1/8loc,t(f),pl}$ .

Guided by the relation (6.1), we define  $V_{s_0^l}^N = V_{s_1^l}^N = \frac{1}{s_1^l - s_0^l} (U_{s_1^l}^N - U_{s_0^l}^N)$  and then:

$$V_{s_{i+1}^l}^N = \frac{2}{s_{i+1}^l - s_i^l} (U_{s_{i+1}^l}^N - U_{s_i^l}^N) - V_{s_i^l}^N.$$

Finally we define the linear interpolation  $V_s^N$ . Note that taking the ultraproduct of defining relations we have

$$(V_{s_0^l}^N)^\omega = (V_{s_1^l}^N)^\omega = \frac{1}{s_1^l - s_0^l} ((U_{s_1^l}^N)^\omega - (U_{s_0^l}^N)^\omega) = \frac{1}{s_1^l - s_0^l} \int_{s_0^l}^{s_1^l} u_s ds = u_{s_0^l}$$

and similarly

$$(V_{s_{i+1}^l}^N)^\omega = \frac{2}{s_{i+1}^l - s_i^l} (\int_{s_i^l}^{s_{i+1}^l} v_s ds) - (V_{s_i^l}^N)^\omega = u_{s_{i+1}^l}.$$

The last equality is by induction on  $i$  using (6.1). Finally by linear interpolation one finds  $(V_s^N)^\omega = u_s$  and  $V_s^N$  are adapted processes and by construction  $\int_0^{s_k^l} V_s^N ds = U_{s_k^l}^N$  so that one deduces the operator norm bound  $\|\int_0^{t_k} V_s^N ds\| \leq C + 1$ . (almost surely).

Moreover, we have a uniform continuity (lipschitzness) in time with value  $L^2$ , uniformly in  $N$  outside of a small interval where we can use a uniform bound in  $L^2$  uniform in  $N$  and thus as in step 2, we have convergence of Riemann integrals (which are indefinite in each  $t_k$ ):

$$\lim_{N \rightarrow \omega} E\left(\frac{1}{2} \int_0^1 \|V_s^N\|_2^2 ds\right) = \frac{1}{2} \int_0^1 \|v_s\|^2 ds, \quad \forall t \in [0, 1], \quad \int_0^t v_s ds = \left(\int_0^t V_s^N ds\right)^\omega.$$

As at the beginning of step 2, one can find  $A_N$  so that  $(W_t^N 1_{A_N})^\omega = S_t$  and  $W_{t_i}^N 1_{A_N}$  is operator norm bounded uniformly in  $N$ . Let us call  $Y_t^N = \int_0^t V_s^N ds + W_t^N 1_{A_N}$  and  $Z_t^N = \int_0^t V_s^N ds + W_t^N$ . Let us call  $R$  the operator norm uniform bound of  $Y_{t_i}^N$ . Since  $(Y_t^N)^\omega = Y_t = \int_0^t v_s ds + S_t$  and since  $W^*(v, Y_{t_1}, \dots, Y_{t_k})$  is a factor, it gives rise to an extremal state in  $\mathcal{S}_R^{mk}$  and thus one can apply proposition 2.10. Indeed, the ultraproduct relation above  $(Y_t^N)^\omega = Y_t$  (and the corresponding relation for their unitary transforms) implies by definition

$$\lim_{N \rightarrow \omega} d(E \circ \tau_{(Y_{t_1}^N, \dots, Y_{t_k}^N, \Upsilon_N)}, \tau_{(Y_{t_1}, \dots, Y_{t_k}, v)}) = 0.$$

Thus, using also (2.5), one gets:

$$\lim_{N \rightarrow \omega} \left| E(G(\tau_{(Y_{t_1}^N, \dots, Y_{t_k}^N, \Upsilon_N)}) - G(\tau_{(Y_{t_1}, \dots, Y_{t_k}, v)})) \right| \leq \lim_{N \rightarrow \omega} C(G) E(d_2(\tau_{(Y_{t_1}^N, \dots, Y_{t_k}^N, \Upsilon_N)}, \tau_{(Y_{t_1}, \dots, Y_{t_k}, v)})) = 0.$$

Moreover, from the uniform bound in lemma 5.3 for (3.3), one deduces as in step 2 (with normalised euclidean norms):

$$\begin{aligned} & \left| E(G(\tau_{(Y_{t_1}^N, \dots, Y_{t_k}^N, \Upsilon_N)}) - G(\tau_{(Z_{t_1}^N, \dots, Z_{t_k}^N, \Upsilon_N)})) \right| \\ & \leq \sqrt{E(\|Y^N - Z^N\|^2)} \times \sqrt{E(3C^3\|Y^N\|_2^2 + 3C^2\|Z^N\|_2^2 + 3D^2)}, \end{aligned}$$

and thus since

$$E(\|Y^N - Z^N\|^2) = E(\|W_t^N\|_2^2 (1_{A_N} - 1)) \rightarrow_{N \rightarrow \infty} 0$$

one gets a limit 0 for our previous expression when  $N \rightarrow \omega$ .

Finally we obtained :

$$\lim_{N \rightarrow \omega} E\left(G(\tau_{(Z_{t_1}^N, \dots, Z_{t_k}^N, \Upsilon_N)}) + \frac{1}{2} \int_0^1 \|V_s^N\|_2^2 ds\right) = G(\tau_{(Y_{t_1}, \dots, Y_{t_k}, v)}) + \frac{1}{2} \int_0^1 \|u_s\|_2^2 ds.$$

But the infimum characterization of theorem 3.3 includes  $V_s^N$  as adapted process and thus

$$E\left(G(\tau_{(Z_{t_1}^N, \dots, Z_{t_k}^N, \Upsilon_N)}) + \frac{1}{2} \int_0^1 \|V_s^N\|_2^2 ds\right) \geq -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}).$$

Taking the limit  $N \rightarrow \omega$  concludes to (6.15), and thus to

$$\lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) = \Lambda_{b, v}(f).$$

Since the limit does not depend on the ultrafilter  $\omega$ , the limit exists as stated.  $\square$

**Corollary 6.3.** Fix  $\Upsilon_N \in \mathcal{U}(M_N(\mathcal{C}))^{\mu\nu}$  (deterministic). Assume that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\Upsilon \in (\mathcal{T}(\mathcal{F}_\mu^\nu), d)$ . **We assume either  $m \geq 2$  or  $m = 1$  and  $W^*(\mu_\Upsilon)$  diffuse.** Let  $\gamma_{sa, N, m} = \gamma_N$  the law of  $m$  hermitian  $N \times N$  brownian motions  $H_s^N \in (M_N(\mathcal{C}))_{sa}^m$ , then, for any  $f \in C_b^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$ , the following limit exists:

$$\lambda(f) := \lim_{N \rightarrow \infty} -\frac{1}{N^2} \log E_{\gamma_{sa, N, m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}).$$

As a consequence,  $\widehat{\sigma}_{\Upsilon_N}^N$  satisfies a Laplace principle in the scale  $N^{-2}$  in  $(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  and thus a Large deviation principle with the good rate function  $I_v : (\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \rightarrow [0, \infty]$ :

$$\begin{aligned} I_v(\tau) &= \sup_{f \in C_b^0(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})} -f(\tau) + \lambda(f) \\ &= \sup_{f \in \mathcal{E}_{reg, \infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})} -f(\tau) + \Lambda_{b,v}(f) \\ &= \sup_{f \in \mathcal{E}_{reg, \infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \cup \mathcal{E}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})} -f(\tau) + \Lambda_{b,v}(f). \end{aligned}$$

*Remark 6.4.* The reader should note the key fact that the limit does not depend on the sequence  $(\Upsilon_N)$  but only of the limiting law  $\mu_\Upsilon$  since  $\Lambda_{b,v}(f)$  depends only of this law. We will use this crucially later in section 8 for our applications to orbital entropy.

*Proof.* Since  $\widehat{\sigma}_{\Upsilon_N}^N$  is exponentially tight by lemma 2.6 and since  $\mathcal{E}_{reg, \infty}(\mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m), d_2)$  is well-separating by lemma 2.3, the first statement is a consequence of [DZ, Th 4.4.10] and the limit obtained in theorem 6.2. The large deviation principle is then a consequence of Bryc's inverse Varadhan lemma [DZ, Th 4.4.2]. The second formula for the good rate function is then contained in [DZ, Ex 4.4.14 p147]. Let us explain the third formula. Consider  $\mathcal{E}$  the smallest class of functions stable by maximum and containing  $\mathcal{E}_{reg, \infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \cup \mathcal{E}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0})$  so that from lemma 2.3, it is well-separating. Since  $\widehat{\sigma}_{\Upsilon_N}^N$  satisfy a large deviation principle, one can apply Varadhan's lemma (see e.g. [DZ, Th 4.3.1]). Since all function in  $\mathcal{E}$  are bounded below, they satisfy the bound (4.3.3) there and thus the exponential transform of theses functionals converge and the hypothesis of [DZ, Th 4.4.2] is satisfied and as before [DZ, Ex 4.4.14 p147] explains why the previous supremum is expressed as a supremum on  $\mathcal{E}$ . From the limit in theorem 6.2, the third expected formula is in between the two suprema (the one on  $\mathcal{E}$  and the one of the second formula), this concludes.  $\square$

We finally conclude the :

*Proof of Theorem C.* The only result not contained in our two previous results is the slightly nicer description of the rate function and its formula for a larger class of functions. Let us explain why  $\Lambda'(v, f) = \Lambda_{b,v}(f)$ , for  $f \in \mathcal{E}_{reg, \infty}$ , first where:

$$\Lambda'(v, f) := \inf \left\{ \frac{1}{2} \int_0^1 \|u_s\|_2^2 ds + f(\tau_{S+\int_0^\cdot u_s ds, v}) : u \in \mathcal{P}'_v \right\}$$

where  $S = (S^1, \dots, S^m)$  are  $m$  standard free brownian motions free from  $v$  in  $(M, \tau)$ , and  $\mathcal{P}'_v$  is the set of control processes in  $L^\infty([0, 1], L^2(M))$  adapted to  $\mathcal{F}_s = W^*(v, S_u, u \leq s)$ , piecewise Hölder continuous of exponent 1/8 (i.e. Hölder continuous on finitely many disjoint open segments with closures covering  $[0, 1]$  but allowing jumps at finitely many of their end point times) and such that, if  $Y_t = S_t + \int_0^t u_s ds$ ,  $Y \in L^\infty([0, 1], M)$  and for any  $t_1, \dots, t_k \in [0, 1]^k$ , there is  $\{s_1, \dots, s_K\} \supset \{t_1, \dots, t_k\}$  with  $W^*(v, Y_{s_1}, \dots, Y_{s_K})$  is a factor.

First the argument in step 2 (iii) in the proof of Theorem 6.2 applies to any set of times containing  $t_1, \dots, t_k$  to get that  $Y_{s_1}, \dots, Y_{s_p}$  has finite free Fisher information relative to  $W^*(v)$  as soon as  $\{s_1, \dots, s_p\} \supset \{t_1, \dots, t_k\}$  and then from [V5] for any set of times (since one can add  $\{t_1, \dots, t_k\}$  and conclude in taking conditional expectations). One concludes to the factoriability of  $W^*(v, Y_{s_1}, \dots, Y_{s_p})$  from [D08] as in the quoted step (for any set of times, which is a stronger condition than above). The uniform operator norm bound on  $Y_s$  is obtained as in step 2 (ii) from 2.7. The process of the step 2 in the proof of Theorem 6.2 thus satisfies all the required properties to get the lower bound with limit  $\Lambda'(v, f)$ . Finally, the same upper bound is enough to conclude to equality since we can replace  $\Lambda_{b,v}(f)$  by  $\Lambda'(v, f)$  in lemma 6.1 (in allowing more times for the points fixed to be  $t_k$ , or said otherwise removing the condition  $s_\omega^i = t_i$ ).

We now extend the classes of  $f$  for which the limit is known to be  $\Lambda'(v, f)$ .

First note that  $\Lambda'(v, \min_{i=1}^n f_i) = \min_{i=1}^n \Lambda'(v, f_i)$  is easy in exchanging a min and an inf. From the proof of [DZ, Th 4.4.10], the equation of the limit extends to  $\min_{i=1}^n f_i$ ,  $f_i \in \mathcal{E}_{reg, \infty}$ .



Using lemma 2.3 and [DZ, lemma 4.4.9], any  $f \in C_b^0(\mathcal{T}(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d)$  is  $\epsilon$  close to  $\min_{i=1}^n f_i$  on a compact set such as  $K_{L\sqrt{\cdot}, 2} \cap \Gamma_L$  from lemma 2.5. Let us show that  $|\Lambda'(v, f) - \Lambda'(v, \min_{i=1}^n f_i)| \leq \epsilon$  if the compact has been chosen as above for  $L$  to be chosen soon.

If  $M$  is a bound for  $f$ , if  $u_s$  is a process almost reaching the infimum for  $\Lambda'(v, f)$  then  $\frac{1}{2} \int \|u_s\|_2^2 \leq M$  (comparing with the value  $u_s = 0$ ) thus  $\|S_t + \int_0^t u_s ds\|_2 \leq 1 + \sqrt{2M}$  and  $\|S_t - S_s + \int_s^t u_v dv\|_2 \leq \sqrt{t-s}(1 + \sqrt{2M})$  implying that for  $Y_t = S_t + \int_0^t u_s ds$ ,  $\tau_Y \in K_{L\sqrt{\cdot}, 2} \cap \Gamma_L$  if  $L = 1 + \sqrt{2M}$ . If we even take  $L = 2 + \sqrt{2M}$  we can do the same for  $\min_{i=1}^n f_i$  since we can replace  $M$  by  $\min_{i=1}^n f_i(\tau_S) \leq M + \epsilon$  since  $\tau_S$  is in the right set. We can thus replace the infimum by the value with the supplementary constraint  $\tau_Y \in K_{L\sqrt{\cdot}, 2} \cap \Gamma_L$  and then it is easy to get the statement  $|\Lambda'(v, f) - \Lambda'(v, \min_{i=1}^n f_i)| \leq \epsilon$ . From the proof of [DZ, Th 4.4.10], one then obtains the limit for  $f \in C_b^0(\mathcal{T}(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d)$ .

Finally, let us show that for any continuous bounded function  $f$ :  $\Lambda'(v, f) = \Lambda(v, f)$ . It suffices to show  $\leq$ . Approximating  $u \in L_{ad}^2([0, 1], L^2(M))$  by  $u \in C_{ad}^0([0, 1], M)$  by standard methods one gets  $\Lambda(v, f)$  gives the same infimum with this condition added and the condition on  $Y_t$  bounded in  $M$  is then fulfilled. Approximating by a piecewise linear function, one gets easily a Lipschitz condition (thus the Hölder continuity assumption). It remains to satisfy the factoriality assumption. One can replace  $u_s$  at small cost by  $v_s$  with  $v_s = 0$  on  $[0, \eta_0]$  and then for  $\eta < \min(\eta_0, t_1)$   $W^*(v, Y_{\eta/2} = S_{\eta/2}, Y_\eta = S_\eta, Y_{t_1}, \dots, Y_{t_k})$  is a factor as in step 2 (viii) of our proof of Theorem 6.2. Thus this approximated process satisfy all the conditions in  $\mathcal{P}'_v$ .  $\square$

We will also need a more technical consequence of the proof in order to compute Voiculescu's microstates free entropy later in some cases.

**Corollary 6.5.** *Let  $f \in \mathcal{E}_{reg,p}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \cup \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}), p \in [2, \infty[, f = G \circ (I_{t_1, \dots, t_k} * Id)$  and consider  $X^{G,N,\Upsilon_N}$  the solution in Theorem 5.4. Then, for any ultrafilter  $\omega \in \beta IN - IN, v = (\Upsilon_N)^\omega, S_t = (W_t^N)^\omega, Y_t = (X_t^{G,N,\Upsilon_N})^\omega \in W^*(v, S_s, s \leq t)$  satisfies an SDE with respect to the canonical brownian motion in  $S_s \in \mathcal{M}_p^\omega$ . There is  $u_s = u_s^G := (b^{G,N,\Upsilon_N}(s, X^{G,N,\Upsilon_N}(s)))^\omega \in L^2(W^*(v, Y_{t_1}, \dots, Y_{t_i}, Y_s))$  for  $t_i < s \leq t_{i+1}$  ( $t_0 = 0, u_s = 0$  if  $s > t_k$ ) such that*

$$Y_t = S_t + \int_0^t u_s^G ds.$$

Finally, the infimum in the definition of  $\Lambda_b(f)$  is reached at  $u_s$ , i.e.:

$$\Lambda_{b,v}(f) = f(\tau_{Y,v}) + \frac{1}{2} \int_0^1 \|u_s^G\|_2^2 ds.$$

*Proof.* All the properties of  $Y_s$  were obtained in step 2 of the proof of Theorem 6.2. For instance, the  $L^2$  space containing  $u_s$  was obtained in (x) based on (ix) (since the drift of  $Z_s = Y_s$  and  $u_s$  coincide). The inequality

$$\lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) \geq f(\tau_{Y,v}) + \frac{1}{2} \int_0^1 \|u_s^G\|_2^2 ds$$

was obtained in (ii), but since from the definition and the upper bound of step 3:

$$f(\tau_{Y,v}) + \frac{1}{2} \int_0^1 \|u_s^G\|_2^2 ds \geq \Lambda_{b,v}(f) \geq \lim_{N \rightarrow \omega} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}).$$

One deduces the stated equality.  $\square$

## 7. APPLICATIONS TO FREE ENTROPY

By the contraction principle of large deviation theory (see e.g. [DZ, Th 4.2.1]) for the projection  $\pi_1 : (\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_\mu^\nu), d_{2,0}) \rightarrow (\mathcal{T}(\mathcal{F}_1^m * \mathcal{F}_\mu^\nu), d)$  at time 1, one deduces

**Theorem 7.1.** Fix  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^\nu$  (deterministic  $\mu = 1$  with previous notation). Assume that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\Upsilon \in (\mathcal{T}(\mathcal{F}_1^\nu), d)$ . **We assume either  $m \geq 2$  or  $m = 1$  and  $W^*(\mu_\Upsilon)$  diffuse.** Then  $\mathfrak{G}_{\Upsilon_N}^N = \pi_1(\widehat{\sigma}_{\Upsilon_N}^N)$  satisfies a Large Deviation Principle in  $(\mathcal{T}(\mathcal{F}_1^m * \mathcal{F}_1^\nu), d)$  with Good rate function  $J_v : (\mathcal{T}(\mathcal{F}_1^m * \mathcal{F}_1^\nu), d) \rightarrow [0, \infty]$  given by

$$J_v(\tau) = \inf \{I_v(\sigma) : \pi_1(\sigma) = \tau, \sigma \in (\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m * \mathcal{F}_1^\nu), d_{2,0})\}.$$

**7.1. Equality of lim inf and lim sup variants of free entropy.** We can now apply this to free entropy as recalled in subsection 2.8. The Large Deviation Principle just shown implies

$$\begin{aligned} \limsup_{N \rightarrow \infty} \left( \frac{1}{N^2} \log (P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{X,v}, \epsilon, K, N))) \right) &\leq - \inf_{\tau \in U_{\epsilon, K}(\tau_{X,v})} J_v(\tau), \\ \liminf_{N \rightarrow \infty} \left( \frac{1}{N^2} \log (P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{X,v}, \epsilon, K, N))) \right) &\geq - \inf_{\tau \in \text{Int}(U_{\epsilon, K}(\tau_{X,v}))} J_v(\tau) \geq - \inf_{\tau \in U_{\epsilon/2, K}(\tau_{X,v})} J_v(\tau). \end{aligned}$$

One thus concludes to our first main application to free entropy, which contains Theorem A.

**Theorem 7.2.** Let  $(M, \tau)$  a finite von Neumann algebra,  $v \in \mathcal{U}(M)^\nu$ ,  $X = (X_1, \dots, X_m)$  self-adjoint elements in  $M$ . Fix  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$ . Assume that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\mu_\Upsilon \in (\mathcal{T}(\mathcal{F}_\mu^\nu), d)$  **and either  $m \geq 2$  or  $m = 1$  and  $W^*(\mu_\Upsilon)$  diffuse.** Then, we have:

$$\begin{aligned} \chi(X|v) &= \underline{\chi}(X|v) = \chi(X|(\Upsilon_N)_{N \in \mathbb{N}}) = \underline{\chi}(X|(\Upsilon_N)_{N \in \mathbb{N}}), \\ \chi^G(X|v) &= \underline{\chi}^G(X|v) = \chi^G(X|(\Upsilon_N)_{N \in \mathbb{N}}) = \underline{\chi}^G(X|(\Upsilon_N)_{N \in \mathbb{N}}), \\ \widetilde{\chi}(\Psi(X)|v) &= \widetilde{\underline{\chi}}(\Psi(X)|v) = -J_v(\tau_{X,v}) = \widetilde{\chi}(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}}) = \widetilde{\underline{\chi}}(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}}). \end{aligned}$$

*Proof.* Fix  $\sigma = \tau_{X,v} \in \mathcal{T}(\mathcal{F}_1^m * \mathcal{F}_\mu^\nu)$ . Since  $-\inf_{\tau \in U_{\epsilon/2, K}(\sigma)} J_v(\tau) \geq -J_v(\sigma)$  and  $J$  is lower semicontinuous so that if  $\tau_{\epsilon, K}$  reaches the infimum (which is reached by compactness since  $J$  good rate function) and  $\epsilon \rightarrow 0, K \rightarrow \infty$  so that  $d(\tau_{\epsilon, K}, \sigma) \rightarrow 0$  one deduces

$$J_v(\sigma) \leq \liminf_{\epsilon, K} J_v(\tau_{\epsilon, K}) = \liminf_{\epsilon, K} \inf_{\tau \in U_{\epsilon/2, K}(\sigma)} J_v(\tau) \leq J_v(\sigma).$$

Thus we have equality and from the previous inequalities obtained from Theorem 7.1, one deduces:

$$-J_v(\tau_{X,v}) = \widetilde{\chi}_\infty(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}}) = \widetilde{\underline{\chi}}_\infty(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}}).$$

Then from Proposition 2.17, one deduces the two last equalities. Note the two obvious relations:

$$\begin{aligned} \widetilde{\chi}(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}}) &\leq \widetilde{\chi}(\Psi(X)|v), \\ \widetilde{\underline{\chi}}(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}}) &\leq \widetilde{\underline{\chi}}(\Psi(X)|v) \leq \widetilde{\chi}(\Psi(X)|v). \end{aligned}$$

This value does not depend on the approximating sequence  $(\Upsilon_N)_{N \in \mathbb{N}}$ , let us choose one giving the two first equalities in the third line. Find first  $\epsilon_n \leq \epsilon_{n-1}/2, K_n > K_{n-1}$  and then  $N'_n > N_n > N'_{n-1}$  such that

$$\sup_{N \in \llbracket N_n, N'_n \rrbracket} \left( \frac{1}{N^2} \sup_{\tau \in V_{\epsilon_n, K_n}(\tau_v)} \log (P(\Gamma_{\infty, \Upsilon}^U(\tau_{\Psi(X), v}, \epsilon_n, K_n, N))) \right) \geq \widetilde{\chi}_\infty(\Psi(X)|v) - \frac{1}{n}.$$

For  $N \in \llbracket N_n, N_{n+1} - 1 \rrbracket$  fix  $\Upsilon_N$  reaching the supremum over the compact set of unitaries. Now fix  $\epsilon, K$ , for  $n$  large such that  $\epsilon > \epsilon_n, K < K_n$ , for  $N \in \llbracket N_n, N_{n+1} - 1 \rrbracket$

$$\begin{aligned} \frac{1}{N^2} \sup_{\tau \in V_{\epsilon_n, K_n}(\tau_v)} \log (P(\Gamma_{\infty, \Upsilon}^U(\tau_{\Psi(X), v}, \epsilon_n, K_n, N))) &= \frac{1}{N^2} \log (P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\Psi(X), v}, \epsilon_n, K_n, N))) \\ &\leq \frac{1}{N^2} \log (P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\Psi(X), v}, \epsilon, K, N))). \end{aligned}$$

Then taking a supremum over  $N \geq N_n$ , one deduces :

$$\begin{aligned} \widetilde{\chi}_\infty(\Psi(X)|v) - \frac{1}{n} &\leq \sup_{N \in \llbracket N_n, N_{n+1} - 1 \rrbracket} \frac{1}{N^2} \log (P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\Psi(X), v}, \epsilon, K, N))) \\ &\leq \sup_{N \geq N_n} \frac{1}{N^2} \log (P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\Psi(X), v}, \epsilon, K, N))) \end{aligned}$$

Then taking the limit  $n \rightarrow \infty, \epsilon \rightarrow 0, K \rightarrow \infty$ , one gets

$$\tilde{\chi}_\infty(\Psi(X)|v) \leq \tilde{\chi}_\infty(\Psi(X)|(\Upsilon_N)_{N \in \mathbb{N}})$$

and this concludes since, of course, by our choice of  $\epsilon_n, K_n, \Upsilon_N$  converges in law to  $v$ .

Finally, from equations (2.17) and (2.16), one deduces all the equalities.  $\square$

We can finally conclude:

*Proof of Theorem A.* One can always assume  $B$  finitely generated (from the definition as an infimum over finitely generated subalgebras of both sides). If  $m \geq 2$  or  $m = 1$  and  $B$  diffuse, this is contained in the previous Theorem 7.2. It remains to consider the case  $m = 1$  and general  $B$ . Consider then  $X_2$  of finite free entropy and free from  $B, X_1$  and use Proposition 2.18:

$$\chi(X_1, X_2|B) = \chi(X_1|B) + \chi(X_2), \quad \underline{\chi}(X_1, X_2|B) = \underline{\chi}(X_1|B) + \chi(X_2).$$

From the case  $m = 2$  we know that  $\chi(X_1, X_2|B) = \underline{\chi}(X_1, X_2|B)$  and since we assumed  $\chi(X_2) > -\infty$ , one deduces the expected  $\chi(X_1|B) = \underline{\chi}(X_1|B)$ .  $\square$

**7.2. Equality  $\chi = \chi^*$  for free Gibbs states with convex potential.** We have to find where the infima in the definition of  $J_v, \Lambda_{b,v}$  and the supremum in the definition of  $I_v$  are reached. We first use the argument in [BCG, Th 7.3].

Let  $\mu \in \mathcal{T}_2(\mathcal{F}_1^m \star \mathcal{F}_1^\nu)$ , we follow [BCG] and define  $\tau_\mu \in \mathcal{T}_2^c(\mathcal{F}_{[0,1]}^m \star \mathcal{F}_1^\nu)$  the law of the brownian bridge, i.e. if  $\{S^1, \dots, S^m\}$  is the law of a free brownian motion free from  $X = \{X^1, \dots, X^m\}, v$  with law (of the unitary  $u(X), v$ )  $\tau_{X,v} = \mu$ , the law  $\tau_{U,v}$  of the process:

$$\left\{ U_t^l = u(tX^l + (1-t)S_{\frac{t}{1-t}}^l), 1 \leq l \leq m, t \in [0, 1] \right\}.$$

**Proposition 7.3.** *Fix the assumption of Theorem 7.1. For any  $\mu \in \mathcal{T}_2(\mathcal{F}_1^m \star \mathcal{F}_1^\nu)$ , we have*

$$J_v(\mu) = I_v(\tau_\mu).$$

*Proof.* Clearly from Theorem 7.1,  $J_v(\mu) \leq I_v(\tau_\mu)$ . But let us recall the straightforward variant of formula (19) in [BCG] for any  $\delta > 0$ :

$$\begin{aligned} \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^2} P(d(\pi_1(\hat{\sigma}_{\Upsilon_N}^N), \mu) \leq \epsilon) \\ = \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^2} P(d(\pi_1(\hat{\sigma}_{\Upsilon_N}^N), \mu) \leq \epsilon, d(\hat{\sigma}_{\Upsilon_N}^N), \tau_\mu) \leq \delta). \end{aligned}$$

Thus, from the large deviation in Corollary 6.3, one deduces (from the fact that  $I$  is a good rate function):

$$\limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^2} P(d(\pi_1(\hat{\sigma}_{\Upsilon_N}^N), \mu) \leq \epsilon) \leq -\lim_{\delta \rightarrow 0} \inf_{d(\tau, \tau_\mu) \leq \delta} I_v(\tau) = -I_v(\tau_\mu)$$

But from the large deviation principle in Theorem 7.1, we also have:

$$\limsup_{\epsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \frac{1}{N^2} P(d(\pi_1(\hat{\sigma}_{\Upsilon_N}^N), \mu) \leq \epsilon) \geq \limsup_{\epsilon \rightarrow 0} -\inf_{d(\tau, \mu) < \epsilon} J_v(\tau) \geq -J_v(\mu).$$

Finally, we thus obtained the missing  $-J_v(\mu) \leq -I_v(\tau_\mu)$ .  $\square$

**Theorem 7.4.** *Fix the assumption of Theorem 7.1. Let  $g \in \mathcal{E}_{app}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_1^m \star \mathcal{F}_1^\nu), d_{2,0})$ . Let  $\tau_g$  the unique solution of  $(SD_g)$  obtained in Theorem 4.4 with fixed law of the unitary part  $v$ , and  $X = X_1, \dots, X_m, v$  having this law. Then, have the equality:*

$$\chi(X_1, \dots, X_m|v) = \chi^*(X_1, \dots, X_m|W^*(v)) \quad \chi^G(X_1, \dots, X_m|v) = \chi^{G*}(X_1, \dots, X_m|W^*(v)).$$

*Proof.* The first equality comes from the second. Let  $\mu_g = \tau_{X,v}$ . By our previous results, corollary 6.3, Theorem 7.2, proposition 7.3, we know that:

$$\begin{aligned} \chi^G(X|v) &= \tilde{\chi}(\Psi(X)|v) = -J(\mu_g) = -I(\tau_{\mu_g}) \\ &= -\sup_{f \in \mathcal{E}_{reg, \infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m \star \mathcal{F}_1^\nu), d_{2,0})} -f(\tau_{\mu_g}) + \Lambda_{b,v}(f). \end{aligned}$$

Thus take  $f \in \mathcal{E}_{reg,\infty}(\mathcal{T}_{2,0}(\mathcal{F}_{[0,1]}^m \star \mathcal{F}_1^\nu), d_{2,0})$ , and fix  $\mathbf{t}(f) = (t_0 = 0 < t_1 < \dots < t_k) \leq 1$ ,  $F \in \mathcal{E}_{reg,\infty}(\mathcal{T}_{2,0}(\mathcal{F}_k^m \star \mathcal{F}_1^\nu), d_{2,0})$  so that  $f = F \circ (I_{t_1, \dots, t_k} * Id)$ . We know from Theorem 6.2 that:

$$\lim_{N \rightarrow \infty} -\frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) = \Lambda_{b,v}(f).$$

If  $t_k = 1$ , call  $\mathbf{t} = \mathbf{t}(f)$  and  $K = k$  and otherwise if  $t_k \neq 1$  define  $\mathbf{t} = (t_0 = 0 < t_1 < \dots < t_k < t_{k+1} = 1)$  and  $K = k + 1$ . Then consider  $G(\tau_{x_1, \dots, x_K, v}) = g(\tau_{x_K, v})$  so that  $G \in \mathcal{E}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_K^m \star \mathcal{F}_1^\nu), d_{2,0})$  then  $\mu_{G,\mathbf{t},N}$  from proposition 2.7 can be seen (after linear change of variable) as a law of the form  $\mu_{g',N}$  as in Theorem 4.4 for  $g' \in \mathcal{E}^{1,1}(\mathcal{T}_{2,0}(\mathcal{F}_1^{Km} \star \mathcal{F}_1^\nu), d_{2,0})$  and thus converges in law, since the law is a marginal of a Hermitian Brownian bridge, it is easy to see (using standard freeness results for instance the characterization of free brownian motion in Theorem 2.20 and concentration from proposition 2.7) the limit law is a marginal of the brownian bridge  $\tau_{\mu_g}$  since  $\mu_{G,\mathbf{t},N}$  is itself the finite dimensional distribution of a brownian bridge (see e.g. [KS] (5.6.28) (5.6.29)) namely of the process  $tX^l + (1-t)H_{\frac{t}{1-t}}^l$  with  $H$  an hermitian brownian motion independent of  $X$ , following the law  $\mu_{g,N}$ . As a consequence, one deduces (using again the concentration result in proposition 2.7 and the lipschitzness of  $F$  with respect to  $d_{2,0}$ ):

$$f(\tau_{\mu_g}) = \lim_{N \rightarrow \infty} E_{\mu_{G,\mathbf{t},N}}(F(\tau, \Upsilon_N)).$$

But of course we can compare this value to

$$\begin{aligned} & -E_{\mu_{G,\mathbf{t},N}}(F(\tau, \Upsilon_N)) - \frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W, \Upsilon_N)}) \\ & \leq \sup_{f \in \mathcal{C}_{sq}(\mathbb{R}^{N^2 m k})} E_{\gamma_{sa,N,m}}(f(\tau_{W_{t_1}, \dots, W_{t_k}}, \Upsilon_N)) \frac{e^{-N^2 g(\tau_{W_1}, \Upsilon_N)}}{Z_{G,\mathbf{t},N}} \\ & \quad - \frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{N^2 f(\tau_{W_{t_1}, \dots, W_{t_k}}, \Upsilon_N)}) \\ & = -\frac{1}{N^2} Ent\left(\frac{1}{Z_{G,\mathbf{t},N}} e^{-N^2 g(\tau_{W_1}, \Upsilon_N)} d\gamma_{sa,N,m} | \gamma_{sa,N,m}\right) \\ & = \mathbf{E} \left( \frac{1}{2} \int_0^1 \|b^{G,N,\Upsilon_N}(t, X^{G,N,\Upsilon_N}(t))\|_2^2 dt \right) \end{aligned}$$

where the next-to-last equality comes from (2.15) and the last one from Theorem 5.4 (with its notation) and its proof. Taking the limit  $N \rightarrow \omega$  and using corollary 6.5 one thus gets:  $-f(\tau_{\mu_g}) + \Lambda_{b,v}(f) \leq \frac{1}{2} \int_0^1 dt \|u_t^G\|_2^2$  and taking the supremum over  $f$ :

$$\chi^G(X_1, \dots, X_m | v) \geq -\frac{1}{2} \int_0^1 dt \|u_t^G\|_2^2.$$

It remains to identify  $u_t^G$ . Note that we know from Theorem 5.4 that the law of  $(X^{G,N,\Upsilon_N}(t), X^{G,N,\Upsilon_N}(1))$  is  $\mu_{g,(t,1),N}$  with the notation of proposition 2.7 and thus the density of  $X^{G,N,\Upsilon_N}(t)$  is the integral on the second variable:

$$\int_{(M_N(\mathbb{C})_{sa})^m} \mu_{g,(t,1),N}(dx_1, dx_2) = \exp \left( -NTr \left( \frac{x_1^2}{2t} \right) - h_t^{N,\Upsilon_N}(\sqrt{N}x_1) \right) dLeb_{(M_N(\mathbb{C})_{sa})^m}(dx_1)$$

We can thus compute the score function of  $X^{G,N,\Upsilon_N}(t)$  to be  $-\frac{N}{t}X^{G,N,\Upsilon_N}(t) - \sqrt{N}\mathcal{D}h_t^{N,\Upsilon_N}(\sqrt{N}X^{G,N,\Upsilon_N}(t))$  and one thus deduces by integration by parts the usual characteristic equation for a non-commutative polynomial  $P \in \mathbb{C}\langle X_1, \dots, X_m, v \rangle$  as in the proof of Theorem 4.4

$$\begin{aligned} & E \left( \frac{1}{N} \text{Tr} \left( \frac{1}{t} X_i^{G,N,\Upsilon_N}(t) P(X^{G,N,\Upsilon_N}(t), \Upsilon_N) \right) \right. \\ & \quad \left. + \frac{1}{N\sqrt{N}} \text{Tr}(\mathcal{D}h_t^{N,\Upsilon_N}(\sqrt{N}X^{G,N,\Upsilon_N}(t)) P(X^{G,N,\Upsilon_N}(t), \Upsilon_N)) \right) \\ & = E \left( \left( \frac{1}{N} \text{Tr} \otimes \frac{1}{N} \text{Tr} \right) (\partial_{X_i} P)(X^{G,N,\Upsilon_N}(t), \Upsilon_N) \right) \end{aligned}$$

and thus in terms of  $b^{G,N,\Upsilon_N}$  this can be written:

$$\begin{aligned} & E \left( \frac{1}{N} \text{Tr} \left( \frac{1}{t} X_i^{G,N,\Upsilon_N}(t) P(X^{G,N,\Upsilon_N}(t), \Upsilon_N) \right) \right. \\ & \quad \left. - \frac{1}{N} \text{Tr}(b_i^{G,N,\Upsilon_N}(t, X^{G,N,\Upsilon_N}(t)) P(X^{G,N,\Upsilon_N}(t), \Upsilon_N)) \right) \\ & = E \left( \left( \frac{1}{N} \text{Tr} \otimes \frac{1}{N} \text{Tr} \right) (\partial_{X_i} P)(X^{G,N,\Upsilon_N}(t), \Upsilon_N) \right) \end{aligned}$$

Note that for further use in the new proof, our previous inequality before the limit can be written if  $-N\Xi^{G,N,\Upsilon_N}(t) = \frac{1}{t}X^{G,N,\Upsilon_N}(t) - b^{G,N,\Upsilon_N}(t, X^{G,N,\Upsilon_N}(t))$  is the score function of  $X^{G,N,\Upsilon_N}(t)$  (from the previous integration by parts formula extended beyond non-commutative polynomials):

$$-E_{\mu_{G,t,N}}(F(\tau, \Upsilon_N)) - \frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_{W,\Upsilon_N})}) \leq \mathbf{E} \left( \frac{1}{2} \int_0^1 \left\| \frac{1}{t} X^{G,N,\Upsilon_N}(t) + N\Xi^{G,N,\Upsilon_N}(t) \right\|_2^2 dt \right)$$

From the concentration result in Theorem 2.7 for  $\mu_{g,(t,1),N}$ , one can take the limit  $N \rightarrow \omega$  of the score function equation and obtain (recall the notation of corollary 6.5  $Y_t = (X^{G,N}(t))^\omega$ )

$$\tau_\omega \left( \frac{Y_t^{(i)}}{t} P(Y_t, v) - (u_t^G)^{(i)} P(Y_t, v) \right) = (\tau_\omega \otimes \tau_\omega)((\partial_{X_i} P)(Y_t, v)).$$

But note the key result in corollary 6.5 that in our situation  $(u_t^G) \in L^2(W^*(Y_t))$  implying that  $\frac{Y_t^{(i)}}{t} - (u_t^G)^{(i)}$  is Voiculescu's  $i$ -th conjugate variable [V5] for  $Y_t$ :  $\xi_i^t$ , and thus our previous inequality reads:

$$\chi^G(X_1, \dots, X_m | v) \geq -\frac{1}{2} \int_0^1 dt \left\| \frac{Y_t}{t} - \xi^t \right\|_2^2 = \chi^{G^*}(X_1, \dots, X_m | W^*(v)).$$

Since the other inequality was known from [BCG] in the case  $W^*(v) = \mathbb{C}$ , one concludes in this case. We give an alternative proof from our result. We use the last formula in Corollary 6.3, one gets:

$$\chi^G(X|v) \leq g(\tau_{\mu_g}) - \Lambda_{b,v}(g) = -\frac{1}{2} \int_0^1 dt \|u_t^G\|_2^2.$$

The last concluding equality comes from corollary 6.5 with the identification of the law  $\tau_{X,v}$  with  $\tau_{\mu_g}$ .  $\square$

**7.3. Proof of Theorem D.** The first equality  $\chi(X_1, \dots, X_m) = \chi^*(X_1, \dots, X_m)$  is in Theorem 7.4. We keep the notation of its proof. We know that  $Y_t$  has the same law as  $tX + (1-t)S_{\frac{t}{1-t}}$ ,  $t \in [0, 1]$  for  $S_t$  a free brownian motion free from  $X = (X_1, \dots, X_m)$  and we know that the conjugate variables are  $\frac{Y_t}{t} - (u_t^G)$ . Most of remaining proof will boil down to a change of time and linear change of variable to induce what we learned in our previous proofs from free brownian bridge to free brownian motion.

**Step 1 :**  $\chi(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m) = \chi^*(X_1 + \sqrt{t}S_1, \dots, X_m + \sqrt{t}S_m).$

The result does not follow from our theorem by lack of a way of producing a universal function  $g_t$  such that the law above satisfies  $(SD_{g_t})$  but the proof will follow closely the one of Theorem 7.4. Since both  $\chi$  (see [V2, Prop 3.6]) and  $\chi^*$  (see [V5, Prop 7.7, 7.8]) have the same formula under linear change of variable, it suffices to prove  $\chi^G(Y_t) = \chi^{G^*}(Y_t)$ . Using [BCG], we are even content to prove :  $\chi^{G^*}(Y_t) \leq \chi^G(Y_t)$ . Let us call  $\mu_t$  the law of  $Y_t$ . Recall that  $\tau_{\mu_t}$  is the law of brownian bridge  $sY_t + (1-s)S'_{\frac{s}{1-s}}$  which is the same law as  $stX + s(1-t)S'_{\frac{t}{1-t}} + (1-s)S'_{\frac{s}{1-s}} = \frac{1}{u} \left( ustX + us(1-t)S'_{\frac{t}{1-t}} + u(1-s)S'_{\frac{s}{1-s}} \right)$  with  $u = u(s,t) = \frac{t}{1-s(1-t)}$ . But since  $u^2[s^2t(1-t) + (1-s)s] = -u^2s^2t^2 + u^2s[st + 1-s] = ust - u^2s^2t^2$  this has the same law as  $\frac{1}{u} \left( ustX + (1-ust)S'_{\frac{ust}{1-ust}} \right)$  which is a time change and linear change of variable of a free Brownian bridge. We call  $L_t(\tau), t \in ]0,1[$  the law of the process  $U_s = \frac{V_{u(s,t)st}}{u(s,t)}, s \in [0,1]$  for  $V_t, t \in [0,1]$  a process of law  $\tau$ . Recall that  $\chi^{G^*}(X_1, \dots, X_m) = -\frac{1}{2} \int_0^1 dt \left\| \frac{Y_t}{t} - \xi^t \right\|_2^2$  where  $\xi^t$  is the conjugate variable for  $Y_t$  thus, since the conjugate variable for  $sY_t + (1-s)S'_{\frac{s}{1-s}}$  is  $u(s,t)\xi^{u(s,t)st}$  one deduces that

$$\begin{aligned} \chi^{G^*}(Y_t) &= -\frac{1}{2} \int_0^1 ds \left\| \frac{Y_{u(s,t)st}}{u(s,t)s} - u(s,t)\xi^{u(s,t)st} \right\|_2^2 \\ &= -\frac{1}{2} \int_0^1 ds \left\| -Y_{u(s,t)st} \left( \frac{1}{t} - 1 \right) - u(s,t) \left( \xi^{u(s,t)st} - \frac{Y_{u(s,t)st}}{u(s,t)st} \right) \right\|_2^2. \end{aligned}$$

Of course we start by the same result as in the proof of our Theorem 7.4:

$$\chi^G(Y_t) = - \sup_{f \in \mathcal{E}_{reg,\infty}(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^m \star \mathcal{F}_1^\nu), d_{2,0})} -f(\tau_{\mu_t}) + \Lambda_{b,v}(f).$$

Fixing  $f$  as in the supremum, and notation similar as in the proof that we follow, we deduce if we call for simplicity  $\mu_{G,N}$  the law of the full brownian bridge process of marginals  $\mu_{G,t,N}$  (in order not to fix the relevant times after the above deterministic change of time):

$$\lim_{N \rightarrow \infty} -E_{\mu_{G,N}}(F(L_t(\tau))) - \frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W)}) = -f(\tau_{\mu_t}) + \Lambda_{b,v}(f).$$

But of course, interpreting the key inequality 7.1 in the proof of Theorem 7.4 in terms of the same computation of the score function as before along an hermitian brownian bridge, on gets:

$$\begin{aligned} &-E_{\mu_{G,N}}(F(L_t(\tau))) - \frac{1}{N^2} \log E_{\gamma_{sa,N,m}}(e^{-N^2 f(\tau_W)}) \\ &\leq \mathbf{E} \left( \frac{1}{2} \int_0^1 \left\| u(s,t)b^{G,N}(u(s,t)st, X^{G,N}(u(s,t)st)) - X^{G,N}(u(s,t)st) \left( \frac{1}{t} - 1 \right) \right\|_2^2 ds \right). \end{aligned}$$

Taking the limit  $N \rightarrow \omega$  after a supremum over  $f$  as in the proof of Theorem 7.4, one thus gets the expected inequality.

**Step 2 :** Use of Time reversal of free brownian motion.

From step 2 (vi) of the proof of Theorem 6.2, we know that  $u_t^G = -\nabla_{Y_t} h_t^\omega(Y_t)$ , which is from (6.6) a  $C$ -Lipschitz function of  $Y_t$ . From the linear change of variable for conjugate variables [V5, Proof of Corol 3.9], in dividing by  $t$  the previous variables,  $Y_t + (t\nabla_{Y_t} h_t^\omega(Y_t)) = t\frac{Y_t}{t} + (t\nabla_{Y_t} h_t^\omega(t\frac{Y_t}{t}))$  is the conjugate variable of  $\frac{Y_t}{t}$  which is of same law as  $X + S_{\frac{1}{t}-1}$  for  $t \leq 1$ .

For  $X \in L^2(W^*(Y_s, s \leq t))^m$ , let us define  $H_t^\omega(X) = t\frac{X^2}{2} + h_t^\omega(tX)$  (using that  $h_t^\omega$  is defined on the same space, even on a huger ultraproduct). Then, the conjugate variable of  $\frac{Y_t}{t}$  is given by  $\nabla_{Y_t/t} H_t^\omega(\frac{Y_t}{t})$ . Note also that from the bounds obtained on  $h_t^\omega$  in step 2.(vi) of the proof of Theorem 6.3, we deduce that  $H_t^\omega$  satisfies (4.5) ( $\alpha = 1/2$ ), (4.6) ( $\beta = 1$ ), the lipschitz, convex and subquadratic behaviour assumed in Theorem 4.3. Note also that from step 2.(ix) of the same proof, we know that  $\nabla_X H_t^\omega(X) \in L^2(W^*(X))$ . This conjugate variable is a  $t + t^2 C$



Lipschitz function of  $\frac{Y_t}{t}$  defined on  $L^2(W^*(Y_s, s \leq t))^m$ . Thus moving to the variable  $s = \frac{1}{t} - 1$ , the brownian motion  $X_s = X + S_s$  has a  $\frac{1}{1+s} + C\frac{1}{(1+s)^2}$  Lipschitz conjugate variable (in the sense it is given by evaluation at  $X_s$  of a Lipschitz map on  $L^2(W^*(X_u, u \geq s))^m$ , given by the gradient of a function  $\mathcal{H}_s^\omega$  with properties similar to  $H_t^\omega$  on this space, thus satisfying the assumptions of Theorem 4.3). Applying, [D14, Prop 15,19], and fixing any  $T \geq s$  the reversed process  $\bar{X}_s = X_{T-s}$  satisfies, for  $\bar{\xi}_s = \xi_{T-s}$  the conjugate variable, the SDE:

$$\bar{X}_t = \bar{X}_0 - \int_0^t \bar{\xi}_u du + \bar{S}_t,$$

for a free brownian motion  $\bar{S}_t$  adapted to the reversed filtration  $L^2(W^*(X_u, u \geq T-t))^m = L^2(W^*(\bar{X}_u, u \leq t))^m$ . We want to apply Theorem 4.3 since  $\bar{\xi}_u = \nabla \mathcal{H}_u^\omega(\bar{X}_u)$ .

Note that from (4.5) ( $\alpha = 1/2$ ), (4.6) ( $\beta = 1$ ), we already have

$$\|\bar{\xi}_s - \bar{\xi}_t\|_2 \leq |t-s|^{1/2}(C + D\|\bar{X}_s\|_2) + \|\bar{X}_s - \bar{X}_t\|_2(C\|\bar{X}_s\|_2 + C\|\bar{X}_t\|_2 + D),$$

and since  $\|\bar{X}_s - \bar{X}_t\|_2 = \|S_{T-s} - S_{T-t}\|_2 = \sqrt{t-s}$  and  $\|\xi_t\|_2$  is bounded, one deduces the stated Hölder continuity of  $\Phi^*(\bar{X}_t)$ . Since  $\bar{\mathcal{F}}_t = L^2(W^*(\bar{X}_0, \bar{S}_u, u \leq t)) \subset L^2(W^*(\bar{X}_u, u \leq t))$ , we can consider the restriction  $\mathcal{H}_u^\omega|_{\bar{\mathcal{F}}_u}$  and apply it Theorem 4.3 to obtain a solution adapted to  $\bar{\mathcal{F}}_u$ :

$$\bar{Z}_t = \bar{X}_0 - \int_0^t P_{\bar{\mathcal{F}}_u}(\nabla \mathcal{H}_u^\omega(\bar{Z}_u)) du + \bar{S}_t.$$

But from the property  $\nabla \mathcal{H}_u^\omega(\bar{Z}_u) \in L^2(W^*(\bar{Z}_u)) \subset \bar{\mathcal{F}}_u$ , it also satisfies:

$$\bar{Z}_t = \bar{X}_0 - \int_0^t \nabla \mathcal{H}_u^\omega(\bar{Z}_u) du + \bar{S}_t.$$

From the uniqueness in Theorem 4.3, this time applied to the unrestricted  $\mathcal{H}_u^\omega$ , we have  $\bar{X}_t = \bar{Z}_t \in W^*(\bar{S}_u, u \leq t, \bar{X}_0)$ . Now as in the proof of corollary 22 in [D14], for  $s \leq T$ :  $\xi_{i,s} - \xi_{i,T} - \int_0^{T-s} \delta_{T-u} \xi_{i,T-u} \# d\bar{S}_u$  is a martingale in the reversed filtration, with increments orthogonal to any stochastic integral adapted to this filtration by proposition 21.(3) in [D14]. But from our adaptedness result to  $W^*(\bar{S}_u, u \leq t, \bar{X}_0)$  and Clarke-Ocone formula [BS], it is such a stochastic integral, thus it is 0 and one deduces in taking the  $L^2$  norm:

$$\|\xi_{i,s}\|_2^2 = \|\xi_{i,T}\|_2^2 + \int_0^{T-s} \|\delta_{T-u} \xi_{i,T-u}\|_2^2 du.$$

Summing over  $i$  gives the stated integral equation.

**Step 3 :** For  $T > 0$ ,  $\chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : \sqrt{T}S_1, \dots, \sqrt{T}S_m) = \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m)$ .

The hardest part is to control entropy in presence which does not seem to be studied at all by usual large deviation techniques. Hopefully our previous construction of a strong solution to the time reversal reduces it to standard results on entropy in presence. It suffices to prove  $\geq$ . Recall that for a subalgebra  $B$ ,

$$\chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : B) = \inf_{n \in \mathbb{N}, Y_1, \dots, Y_n \in B} \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : Y_1, \dots, Y_n).$$

From [V1, Prop 3.4], since  $B = W^*(\bar{S}_t, t \leq T)$  is free from  $X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m$ , one deduces:

$$\begin{aligned} \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m) &= \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : B) \\ &= \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m, B) \\ &= \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m, B, \sqrt{T}S_1, \dots, \sqrt{T}S_m) \\ &\leq \chi(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m : \sqrt{T}S_1, \dots, \sqrt{T}S_m). \end{aligned}$$

Indeed the second line comes from [V2, Prop 1.7] and the third line comes from the fact that  $\sqrt{T}S_i = \overline{X}_0 - \overline{X}_T \in W^*(X_1 + \sqrt{T}S_1, \dots, X_m + \sqrt{T}S_m, B)$  from our result in step 2 and from [V2, Corol 1.8] (slightly extended to the non-finitely algebra case). The last inequality is then a trivial and concluding inequality.

## 8. APPLICATIONS TO HAAR UNITARIES AND ORBITAL ENTROPY

**8.1. Inverse contraction and contraction giving LDP for Haar unitaries.** Consider  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))^{m\nu}$  is a bunch of (deterministic) unitary matrices and as before

$$\widehat{\sigma}_\Upsilon^N = \tau_{H^N, \Upsilon} \in \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{F}_\nu^m).$$

Note that we now consider  $2m$  variables instead of  $m$  in order to associate complex gaussians matrices  $C_i^N = \frac{H_{i,t=1}^N + \sqrt{-1}H_{m+i,t=1}^N}{\sqrt{2}}, i \leq m$ . Then from [AGZ, Exercice 5.4.17] or [HP, p 164], the unitary in the polar decomposition  $U_i^N = C_i^N((C_i^N)^* C_i^N + 1/n)^{-1/2}$  follows the Haar measure on  $\mathcal{U}(M_N(\mathbb{C}))$  (especially it is almost surely unitary and not only a partial isometry). We can thus consider

$$\widehat{\Sigma}_\Upsilon^N = \tau_{U^N, \Upsilon} \in \mathcal{T}(\mathcal{F}_{\nu+1}^m).$$

To obtain a large deviation principle for  $\widehat{\Sigma}_\Upsilon^N$ , we follow the ideas in [CDG1, section 4.5].

Unfortunately, it is not clear that there is an associated continuous map. To obtain such a map, it is more natural to consider  $V_i^{N,n} = C_i^N((C_i^N)^* C_i^N + 1/n)^{-1/2}$  which is not unitary though. Thus consider  $\mathcal{G}^m * \mathcal{F}_\nu^m$  the universal  $C^*$ -algebra generated by  $m$  contractions  $V_i$  with  $\|V_i\| \leq 1$  and  $m\nu$  unitaries. Replacing inverses by adjoints in the definition of  $d$  on  $\mathcal{F}_{\nu+1}^m$  one obtains an isometric inclusion of tracial state spaces:

$$(\mathcal{T}(\mathcal{F}_{\nu+1}^m), d) \subset (\mathcal{T}(\mathcal{G}^m * \mathcal{F}_\nu^m), d)$$

and  $d$  obviously gives a metric for the weak-\* topology. (Note also that with the definition bellow  $d(\tau_{V,v}, \tau_{V',v'}) = d_{0,0}(\tau_{1,V,1,v}, \tau_{1,V',1,v'})$ ).

Similarly, one can consider  $\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m$  with a unitary process added and a subset of the tracial state with moments and continuity conditions as in subsection 2.1:  $\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m)$ . We now define explicitly the metric  $d_{2,0}$  on this space similarly. To fix a convenient notation for moments we call  $v_t^i = u_t^i, i \leq m, t \in [0, 1]$  the  $m$  first unitary process variables,  $v_{2+t}^i = u_t^{m+i}, i \leq m, t \in [0, 1]$ , the  $m$  second process variables,  $v_j^i$  a  $j = 4, 5$   $i = 1, \dots, m$  the contraction variables of  $\mathcal{G}^{2m}$  and  $v_j^i$  a  $j = 6, \dots, 5 + \nu$   $i = 1, \dots, m$  the last  $\nu m$  unitary variables.

$$\begin{aligned} d_{2,0}(\tau_1, \tau_2) &= d_{0,0}(\tau_1, \tau_2) + \sup_{l=1, \dots, 2m} \sup_{t \in [0,1]} \left| (\tau_1 - \tau_2) \left( \left( \frac{u_t^l + 1}{u_t^l - 1} \right)^* \left( \frac{u_t^l + 1}{u_t^l - 1} \right) \right) \right| \\ &+ \sum_{k=1}^{\infty} 2^{-k} \sup_{t \in [0,1]} \sup_{(l, i_1, \dots, i_m) \in \llbracket 1, 2m \rrbracket \times \llbracket 1, m \rrbracket^k} \\ &\sup_{(\epsilon_1, \dots, \epsilon_m) \in \{*, 1\}^k} \sup_{(j_1, \dots, j_k) \in (\{4, \dots, 5 + \nu\} \cup [0, 1] \cup [2, 3])^k} |(\tau_1 - \tau_2) \left( \left( \frac{u_t^l + 1}{u_t^l - 1} \right)^* (v_{j_1}^{i_1})^{\epsilon_1} \dots (v_{j_k}^{i_k})^{\epsilon_k} \right)|. \end{aligned}$$

where

$$\begin{aligned} d_{0,0}(\tau_1, \tau_2) &= \sum_{k=1}^{\infty} 2^{-k} \\ &\sup_{(i_1, \dots, i_m) \in \llbracket 1, m \rrbracket^k} \sup_{(\epsilon_1, \dots, \epsilon_m) \in \{*, 1\}^k} \sup_{(j_1, \dots, j_k) \in \{4, \dots, 5 + \nu\} \cup [0, 1] \cup [2, 3]^k} |(\tau_1 - \tau_2) \left( (v_{j_1}^{i_1})^{\epsilon_1} \dots (v_{j_k}^{i_k})^{\epsilon_k} \right)|. \end{aligned}$$

If we consider also  $\mathcal{V}_i^{N,n} = ((C_i^N)^* C_i^N + 1/n)^{-1}$ , we obtain two laws :

$$\widehat{\Sigma}_\Upsilon^{N,n} = \tau_{V^{N,n}, \Upsilon} \in \mathcal{T}(\mathcal{G}^m * \mathcal{F}_\nu^m), \quad \widehat{\Theta}_\Upsilon^{N,n} = \tau_{H^N, V^{N,n}, \mathcal{V}^{N,n}, \Upsilon} \in \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m).$$

Consider the continuous projections  $g^n : \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m) \rightarrow \mathcal{T}(\mathcal{G}^m * \mathcal{F}_1^\nu)$  forgetting the law of middle  $m$  second contraction variables and of the unitary processes and  $j^n : \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m) \rightarrow \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{F}_\nu^m)$  forgetting the law of middle  $2m$  contraction variables.

Consider finally  $i^n : \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{F}_1^\nu) \rightarrow \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m)$  the canonical map associating to the law of  $\tau = \tau_{X,v}$  the law  $\tau_{X,V^n,\mathcal{V}^n,v}$  with the variables  $V^n, \mathcal{V}^n$  associated in the GNS representation of  $\tau$ ,

$$V_i^n = C_i((C_i)^* C_i + 1/n)^{-1/2}, \mathcal{V}_i^n = ((C_i)^* C_i + 1/n)^{-1}, C_i = \frac{4}{\sqrt{2}} \left( \sqrt{-1} \frac{u_1^i + 1}{u_1^i - 1} - \frac{u_1^{i+m} + 1}{u_1^{i+m} - 1} \right)$$

Thus we have  $j^n \circ i_n = id$  and the relations:

$$g^n(\widehat{\mathfrak{S}}_\Upsilon^{N,n}) = \widehat{\Sigma}_\Upsilon^{N,n}, \quad j^n(\widehat{\mathfrak{S}}_\Upsilon^{N,n}) = \widehat{\sigma}_\Upsilon^N, \quad i^n(\widehat{\sigma}_\Upsilon^N) = \widehat{\mathfrak{S}}_\Upsilon^{N,n}.$$

We aim at using the inverse contraction principle [DZ, Th 4.2.4] to  $j^n$  and then the contraction principle to  $g^n$ . We gather the byproduct in the next:

**Lemma 8.1.** *Consider  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))^{m\nu}$  is a bunch of (deterministic) unitary matrices such that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\tau_\nu \in (\mathcal{T}(\mathcal{F}_\nu^m), d)$ . With  $m \geq 1$  and the previous notation,  $\widehat{\mathfrak{S}}_\Upsilon^{N,n}$  satisfy a large deviation principle in the scale  $N^{-2}$  in  $i^n \left( \mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{F}_1^\nu) \right) \subset (\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m, d_{2,0}))$  with good rate function  $I_\nu \circ j^n$ .*

Moreover,  $\widehat{\Sigma}_\Upsilon^{N,n}$  satisfy a large deviation principle in the scale  $N^{-2}$  in  $(\mathcal{T}(\mathcal{G}^m * \mathcal{F}_1^\nu), d)$  with good rate function

$$I'_{\nu,n}(\mu) = \inf \{ I_\nu \circ j^n(\tau) : g^n(\tau) = \mu \}.$$

*Remark 8.2.* Note that with the contraction principle, this result also implies a large deviation principle for  $m$  complex Wishart matrices  $(C_i^N)^* C_i^N$ . This imply the usual result for one Wishart matrix since bounded continuous functions of  $(C_i^N)^* C_i^N$  and  $((C_i^N)^* C_i^N + 1)^{-1}$  coincide. We could also treat similarly more general covariances and non-centered cases  $(C_i^N - A_N^i)^* B_N^i (C_i^N - A_N^i)$  as soon as  $B_N^i \geq 0, A_N^i$  are deterministic and uniformly bounded and converge jointly with  $\Upsilon_N$  to a non-commutative distribution. Indeed,  $A_N, B_N$  can be expressed as a sum of unitary matrices (by functional calculus) with convergent joint law to which we can apply our first LDP.

*Proof.* The second LDP is a direct consequence of the first and the contraction principle (see e.g. [DZ, Th 4.2.1]) since  $g^n$  is continuous. Similarly  $j^n$  is continuous on  $Im(i_n)$  and it is a bijection of inverse  $i^n$ . In order to apply [DZ, Th 4.2.4], it remains to check that  $\widehat{\mathfrak{S}}_\Upsilon^{N,n}$  is exponentially tight in  $Im(i_n)$ . Of course, since  $i_n$  is a bijection (onto its image)  $\widehat{\mathfrak{S}}_\Upsilon^{N,n} \in i_n((K_{L\sqrt{\epsilon},2} \cap \Gamma_L))$  if and only if  $\widehat{\sigma}_\Upsilon^N \in (K_{L\sqrt{\epsilon},2} \cap \Gamma_L)$ . Thus from Lemma 2.6, it suffices to check that  $i_n((K_{L\sqrt{\epsilon},2} \cap \Gamma_L))$  is compact for  $d_{2,0}$ . From an argument similar to lemma 2.5, it is easy to see that any sequence has a subsequence converging in  $(\mathcal{T}_{2,0}^c(\mathcal{F}_{[0,1]}^{2m} * \mathcal{G}^{2m} * \mathcal{F}_\nu^m, d_{2,0}))$ . It remains to check that  $Im(i_n)$  is a closed subset, thus take a sequence  $\tau_N$  converging to  $\tau$ . Since  $\|\mathcal{V}_i^n\| \leq n$  and  $\mathcal{V}_i^n \geq 0$ , for any  $\epsilon$ , by functional calculus and Stone-Weierstrass Theorem, there is a polynomial such that  $\|P_\epsilon(\mathcal{V}_i^n) - (\mathcal{V}_i^n)^{1/2}\| \leq \epsilon$ .

Then any state in  $Im(i_n)$  (such as  $\tau_N$ ) satisfies a universal bound  $|\tau_N([C_i^* P_\epsilon(\mathcal{V}_i^n) - V_i^n]U)| \leq \tau_N(C_i^* C_i)^{1/2} \epsilon$  for  $U$  any product of unitaries and contractions appearing in the definition of  $\tau_N$ , thus passing to the limit,  $\tau$  satisfies the same equation, which implies  $C_i(\mathcal{V}_i^n)^{1/2} = V_i^n$  in the GNS representation as it should. Similarly we have

$$|\tau_N([C_i^* V_i^n P_\epsilon(\mathcal{V}_i^n)U + \frac{1}{n} \mathcal{V}_i^n - 1]U)| \leq \tau_N(C_i^* C_i)^{1/2} \epsilon$$

which implies in the GNS representation of the limit state  $\tau$  that  $(C_i^* C_i + 1/n) \mathcal{V}_i^n = 1$  and since both sides are self-adjoint, in taking the adjoint, one gets  $\mathcal{V}_i^n = (C_i^* C_i + 1/n)^{-1}$ . This last relation guaranties  $\tau \in Im(i_n)$ .  $\square$

We are now ready to obtain the main result of this section:

**Theorem 8.3.** Consider  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))^{m\nu}$  is a bunch of (deterministic) unitary matrices such that the non-commutative law  $\tau_{\Upsilon_N}$  converges to some  $\tau_v \in (\mathcal{T}(\mathcal{F}_\nu^m), d)$ . With  $m \geq 1$  and the previous notation,  $\widehat{\Sigma}_\Upsilon^N$  satisfy a large deviation principle in the scale  $N^{-2}$  in  $(\mathcal{T}(\mathcal{G}^m * \mathcal{F}_1^\nu), d)$  with (good) rate function:

$$I'_{\tau_v}(\mu) = \sup_{\delta > 0} \liminf_{n \rightarrow \infty} \inf \{I'_{v,n}(\tau) : d(\tau, \mu) < \delta\}.$$

*Proof.* Note that for any  $C_j$  such that  $C_j(C_j^*C_j)^{-1/2}$  is unitary (e.g.  $C_j^N$ ), we have:

$$\begin{aligned} C_j(C_j^*C_j + \epsilon)^{-1/2} - C_j(C_j^*C_j)^{-1/2} &= C_j \left( (C_j^*C_j + \epsilon)^{1/2} - (C_j^*C_j)^{1/2} \right) (C_j^*C_j + \epsilon)^{-1} \\ &\quad - C_j(C_j^*C_j)^{-1/2} \epsilon (C_j^*C_j + \epsilon)^{-1} \\ &= \frac{1}{2} C_j \left( \int_0^\epsilon (C_j^*C_j + u)^{-1/2} \right) (C_j^*C_j + \epsilon)^{-1} \\ &\quad - C_j(C_j^*C_j)^{-1/2} \epsilon (C_j^*C_j + \epsilon)^{-1} \end{aligned}$$

Knowing that  $(C_j^*C_j + u)^{-1/2} C_j^*C_j (C_j^*C_j + u)^{-1/2} \leq 1$ , one deduces in taking the  $L^1$  norm that

$$\|C_j(C_j^*C_j)^{-1/2} - C_j(C_j^*C_j + \epsilon)^{-1/2}\|_1 \leq \frac{3\epsilon}{2} \|(C_j^*C_j + \epsilon)^{-1}\|_1.$$

From that, one deduces:

$$d(\widehat{\Sigma}_\Upsilon^N, \widehat{\Sigma}_\Upsilon^{N,n}) \leq 3 \sum_{j=1}^m \frac{1}{N} \text{Tr} \left( \frac{n^{-1}}{(C_j^N)^* C_j^N + n^{-1}} \right).$$

Thus, our next lemma implies  $\widehat{\Sigma}_\Upsilon^{N,n}$  are exponentially good approximations of  $\widehat{\Sigma}_\Upsilon^N$  (see [DZ, Def 4.2.14]). Then, since our previous lemma showed a large deviation principle for  $\widehat{\Sigma}_\Upsilon^{N,n}$  with good rate functions, one deduces: from [DZ, Th 4.2.16] that  $\widehat{\Sigma}_\Upsilon^N$  satisfies a weak LDP with the stated rate function, but since the target space is compact (in the weak-\* topology given by the metric  $d$ ), this is also a full LDP with good rate function.  $\square$

It only remains to get the missing exponential tightness result which follows the idea in [CDG1, (4.23)].

**Lemma 8.4.** With the previous notation, for every  $\epsilon \in ]0, 1[$ , we have :

$$\limsup_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N^2} \log P \left( \sum_{j=1}^m \frac{1}{N} \text{Tr} \left( \frac{n^{-1}}{(C_j^N)^* C_j^N + n^{-1}} \right) \geq \epsilon \right) = -\infty.$$

*Proof.* Since each  $C_i^*C_i$  is a complex Wishart matrix, we will use a standard large deviation result for them (see e.g. [HP, Th 5.5.1], and also the remark after (5.5.7) for the identification of the rate function in the complex case and [HP, Th 5.5.7] for the computation of the constant.) Namely, we use that the law (spectral distribution) of  $C_i^*C_i$  under law  $\widehat{\sigma}_\Upsilon^N$  satisfies a large deviation principle in the scale  $N^{-2}$  in the space  $\mathcal{P}(\mathbb{R}^+)$  of probability measures on  $\mathbb{R}^+$  with good rate function

$$I_1(\mu) = -\Sigma(\mu) + \int x d\mu(x) - \frac{3}{2},$$

with  $\Sigma(\mu) = \int \int \log|x-y| d\mu(x) d\mu(y)$ . We follow [CDG1, p1258] Define  $A_\eta^\epsilon = \{\mu \in \mathcal{P}(\mathbb{R}_+) : \int d\mu(x) \frac{\eta}{x+\eta} \geq \epsilon\}$  for  $\eta, \epsilon > 0$ . Then, from the above stated LDP, we have:

$$\limsup_{N \rightarrow \infty} \frac{1}{N^2} \log P \left( \widehat{\sigma}_\Upsilon^N \left( \frac{\eta}{C_i^*C_i + \eta} \right) \geq \epsilon \right) \leq -\inf \{I_1(\mu), \mu \in A_\eta^\epsilon\}.$$

Arguing as in [CDG1, p1258], one gets two universal finite constants  $c > 0, c'$  such that for  $\eta \leq \eta_0$  small enough (independent of  $\epsilon \in ]0, 1[$ ), for all  $\mu \in A_\eta^\epsilon$ ,  $I_1(\mu) \geq c(|\log(\eta)|\epsilon^2 + c')$ . Thus, if

$F(\eta, \epsilon) = -\inf\{I_1(\mu), \mu \in A_\eta^\epsilon\}$  one gets

$$\sup_{0 < \eta \leq \eta_0} -\inf\{I_1(\mu), \mu \in A_\eta^\epsilon\} \leq -c(|\log(\eta)|\epsilon^2 + c').$$

This concludes in taking  $\eta = n^{-1}$ . □

**8.2. Equivalent definitions for orbital entropy and its general properties.** We first recall (an obvious variant using unitary variables instead of self-adjoint ones of) the definition from [Ue14]. We won't use the variants we introduced in [BD13] using random microstates that would not be adapted to our current large deviation setting but we introduce yet another variant too. In the case of hyperfinite multi-variables considered in [HMU], the new variant also coincides with the original definition but it is better behaved for non-hyperfinite variables.

Let  $\mathbf{U}_i = (U_{i1}, \dots, U_{i\nu}), v = (v_1^1, \dots, v_\nu^1, \dots, v_\nu^\mu), 1 \leq i \leq m$ , be arbitrary random multi-variables in  $(\mathcal{M}, \tau)$  (i.e. finite tuples of unitaries). Let  $K \in \mathbb{N}$  and  $\delta > 0$  be given and  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))^{\nu\mu}$ . Recall we defined in subsection 2.8 the set of matricial microstates

$\Gamma_{\infty, \Upsilon}^U(\tau_{\mathbf{U}_1 \sqcup \dots \sqcup \mathbf{U}_m, v}; \delta, K, N)$ . Consider the map  $\Phi_N : \mathcal{U}(M_N(\mathbb{C}))^m \times \left(\prod_{i=1}^{m+\mu} (\mathcal{U}(M_N(\mathbb{C})))^\nu\right) \rightarrow \left(\prod_{i=1}^{m+\mu} (\mathcal{U}(M_N(\mathbb{C})))^\nu\right)$  (that we may use sometimes with  $\mu = 0$ ) defined by :

$$\Phi_N((U_i)_{i=1}^m, (\mathbf{V}_i)_{i=1}^{m+\mu}) := (U_1 \mathbf{V}_1 U_1^*, \dots, U_m \mathbf{V}_m U_m^*, \mathbf{V}_{m+1}, \dots, \mathbf{V}_{m+\mu})$$

with  $U_i \mathbf{V}_i U_i^* := (U_i V_{i1} U_i^*, \dots, U_i V_{i\nu} U_i^*)$ . Recall also that for given multi-matrices  $\mathbf{V}_i = (V_{ij})_{j=1}^\nu \in (\mathcal{U}(M_N(\mathbb{C})))^\nu, 1 \leq i \leq n$ , the set of orbital microstates relative to  $\Upsilon$  is

$$\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : (\mathbf{V}_i)_{i=1}^m | \Upsilon; \delta, K, N) \equiv \Gamma_{\text{orb}}(\tau_{\mathbf{U}_1, \dots, \mathbf{U}_m, v} : (\mathbf{V}_i)_{i=1}^m | \Upsilon; \delta, K, N)$$

is defined to be all  $(U_i)_{i=1}^m \in \mathcal{U}(M_N(\mathbb{C}))^m$  such that  $\Phi_N((U_i)_{i=1}^m, (\mathbf{V}_i)_{i=1}^m, \Upsilon) \in \Gamma_{\infty, \Upsilon}^U(\tau_{\mathbf{U}_1 \sqcup \dots \sqcup \mathbf{U}_m, v}; \delta, K, N)$ .

We also call  $\gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m}$  the tensor product of Haar measures.

**Definition 8.5** (Ueda's definition in Prop 2.4 of [Ue14]). We define

$$(8.1) \quad \bar{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta, K, N) = \sup_{\mathbf{V}_i \in \Gamma_{\infty}^U(\mathbf{U}_i; \delta, K, N)} \sup_{\tau_{\Upsilon} \in V_{\epsilon, K}(\tau_v)} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} (\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : (\mathbf{V}_i)_{i=1}^m | \Upsilon; \delta, K, N)) \right)$$

(defined to be  $-\infty$  if the set is empty) and

$$(8.2) \quad \bar{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) = \lim_{\substack{K \rightarrow \infty \\ \delta \searrow 0}} \limsup_{N \rightarrow \infty} \frac{1}{N^2} \bar{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta, K, N).$$

Recall that  $\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v$  is the free product state. Our slightly different definition comes from the idea that to obtain from our results an orbital entropy that could be defined as in [HMU] using a sequence, we need to fix the full joint law, and the most natural choice for that since orbital entropy is supposed to be an entropy relative to the free product state is to take of course the free product state as reference joint state. This gives the following definition:

**Definition 8.6.** We write for  $\delta > 0, K, N \in \mathbb{N}$  and  $\epsilon = (\epsilon_n)_{n \in \mathbb{N}}$  a non-increasing positive sequence with  $\epsilon_n \rightarrow 0$  and we call their set *Seq*.

$$(8.3) \quad \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta, K, N) = \sup_{(\mathbf{V}_1, \dots, \mathbf{V}_m, \Upsilon) \in \Gamma_{\infty, \emptyset}^U(\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v; \delta, K, N)} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} (\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : (\mathbf{V}_i)_{i=1}^m | \Upsilon; \delta, K, N)) \right)$$

(defined to be  $-\infty$  if the set is empty).

$$(8.4) \quad \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta, K, N; \epsilon) = \inf_{(\mathbf{V}_1, \dots, \mathbf{V}_m, \Upsilon) \in \Gamma_{\infty, \emptyset}^U(\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v; \epsilon_N, \lfloor \epsilon_N^{-1} \rfloor, N)} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} (\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : (\mathbf{V}_i)_{i=1}^m | \Upsilon; \delta, K, N)) \right)$$

(defined to be 0 if the set on which the infimum is taken is empty and  $-\infty$  if the set under the volume is empty). We define the *orbital free entropy of  $\mathbf{U}_1, \dots, \mathbf{U}_n$  relative to  $v$*  to be:

$$(8.5) \quad \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \lim_{\substack{K \rightarrow \infty \\ \delta \searrow 0}} \limsup_{N \rightarrow \infty} \frac{1}{N^2} \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v; \delta, K, N).$$

For von Neumann algebras  $A_1, \dots, A_m, B$  we define the *orbital free entropy of  $A_1, \dots, A_m$  relative to  $B$*  to be:

$$\chi_{\text{orb}}(A_1, \dots, A_m|B) = \inf_{\mathbf{U}_1 \subset A_1} \dots \inf_{\mathbf{U}_m \subset A_m} \inf_{v \subset B} \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v).$$

We also define temporarily the variants:

$$(8.6) \quad \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \lim_{\substack{K \rightarrow \infty \\ \delta \searrow 0}} \liminf_{N \rightarrow \infty} \frac{1}{N^2} \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v; \delta, K, N),$$

$$(8.7) \quad \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v; \epsilon) = \lim_{\substack{K \rightarrow \infty \\ \delta \searrow 0}} \liminf_{N \rightarrow \infty} \frac{1}{N^2} \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v; \delta, K, N; \epsilon),$$

$$\underline{\underline{\chi}}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \inf_{\epsilon \in \text{Seq}} \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v; \epsilon)$$

In the spirit of [HMU], one can consider  $\Xi_N = (\mathbf{V}_1^N, \dots, \mathbf{V}_m^N) \in \mathcal{U}(M_N(\mathbb{C}))^{m\nu}$  a sequence of multivariables with  $\tau_{\Xi_N, \Upsilon_N}$  converging in  $(\mathcal{T}(\mathcal{F}_\nu^m * \mathcal{F}_\nu^\mu), d)$  to  $\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v$ . We then define:

$$(8.8) \quad \begin{aligned} & \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \\ &= \lim_{\substack{K \rightarrow \infty \\ \delta \searrow 0}} \limsup_{N \rightarrow \infty} \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m}(\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N)) \right) \end{aligned}$$

$$(8.9) \quad \begin{aligned} & \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \\ &= \lim_{\substack{K \rightarrow \infty \\ \delta \searrow 0}} \liminf_{N \rightarrow \infty} \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m}(\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N)) \right) \end{aligned}$$

Our main result concerning orbital free entropy is as follows:

**Theorem 8.7.** *Let  $\Xi_N = (\mathbf{V}_1^N, \dots, \mathbf{V}_m^N) \in \mathcal{U}(M_N(\mathbb{C}))^{m\nu}, \Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))^{\mu\nu}$  sequences of multivariables with  $\tau_{\Xi_N, \Upsilon_N}$  converging in  $(\mathcal{T}(\mathcal{F}_\nu^m * \mathcal{F}_\nu^\mu), d)$  to  $\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v$ . Then we have the (in)equalities:*

$$\begin{aligned} & \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) = \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \\ &= \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m, v) \leq \bar{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) \end{aligned}$$

Moreover we have  $\underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m, v)$  if  $\mathbf{U}_1, \dots, \mathbf{U}_m, v$  have each one separately finite dimensional approximants. Especially, all those definitions coincide with the one of [HMU] when  $\mu = 0$  and  $W^*(\mathbf{U}_i)$  are hyperfinite.

*Proof.* The inequality  $\chi_{\text{orb}} \leq \bar{\chi}_{\text{orb}}$  is obvious from the definitions since one set where the supremum is taken is smaller. The equality  $\chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m|v) = \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m, v)$  is elementary since

$$\begin{aligned} & \Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m, v : ((\mathbf{V}_i)_{i=1}^m, \Upsilon); \delta, K, N) \\ &= \{ (U_1, \dots, U_{m+1}) : (U_1 U_{m+1}^{-1}, \dots, U_m U_{m+1}^{-1}) \in \Gamma_{\text{orb}}(\tau_{\mathbf{U}_1, \dots, \mathbf{U}_m, v} : (\mathbf{V}_i)_{i=1}^m | \Upsilon; \delta, K, N) \}, \end{aligned}$$

and by unitary invariance of Haar measure, this implies that the volume considered in the definition of both entropies is the same.

We take  $(\Xi, \Upsilon)$  as the sequence of deterministic matrices in Theorem 8.3, to conclude that  $\hat{\Sigma}_{(\Xi, \Upsilon)}^N$  satisfy a large deviation principle in  $(\mathcal{T}(\mathcal{F}_1^m * \mathcal{F}^{m+\mu\nu}), d)$  with good rate function  $I'_{\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v}$  which does not depend on the sequence  $(\Xi, \Upsilon)$  but only of its limit law.



We consider the continuous map  $\Phi : (\mathcal{T}(\mathcal{F}_1^m * \mathcal{F}_\nu^{m+\mu}), d) \rightarrow (\mathcal{T}(\mathcal{F}_\nu^m * \mathcal{F}_\nu^m), d)$  given by

$$\Phi(\tau_{U_1, \dots, U_m, \mathbf{U}_1, \dots, \mathbf{U}_m, v}) = \tau_{U_1 \mathbf{U}_1 U_1^*, \dots, U_m \mathbf{U}_m U_m^*, v}.$$

Note that it corresponds to  $\Phi_N$  at the level of states  $\tau_{\Phi_N((U_i)_{i=1}^n, (\mathbf{V}_i)_{i=1}^{m+\mu})} = \Phi(\tau_{(U_i)_{i=1}^n, (\mathbf{V}_i)_{i=1}^{m+\mu}})$ .

If  $U_1^N, \dots, U_m^N$  are independent Haar random unitaries, one sees that  $\Phi(\widehat{\Sigma}_{(\Xi, \Upsilon)}^N) = \tau_{\Phi_N((U_i^N)_{i=1}^m, (\Xi_N, \Upsilon_N))}$  satisfies a large deviation principle in the same scale  $N^{-2}$  by the contraction principle, with good rate function

$$J'_{\tau_{U_1} * \dots * \tau_{U_m} * \tau_v}(\mu) = \inf\{I'_{\tau_{U_1} * \dots * \tau_{U_m} * \tau_v}(\tau) : \Phi(\tau) = \mu\}.$$

One thus deduces as in the proof of Theorem 7.2 that

$$\begin{aligned} \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) &= \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \\ &= J'_{\tau_{U_1} * \dots * \tau_{U_m} * \tau_v}(\tau_{\mathbf{U}_1, \dots, \mathbf{U}_m, v}). \end{aligned}$$

And the result is thus independent of the sequences  $((\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}})$  converging in law to the free product trace  $\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v$ . To conclude with the other two equalities, it suffices to build appropriate sequences with limit equal to  $\underline{\chi}_{\text{orb}}$  and  $\chi_{\text{orb}}$ . Again, the argument for  $\chi_{\text{orb}}, \underline{\chi}_{\text{orb}}$  are similar to the one in Theorem 7.2, we detail only the case of  $\underline{\chi}_{\text{orb}}$ . Note first that we always have :

$$\underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \epsilon) \leq \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}})$$

as soon as  $d(\tau_{\Xi_N, \Upsilon_N}, \tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v) < \epsilon_N 2^{-\lfloor \epsilon_N^{-1} \rfloor}$  for all  $N$  and there is of course such a sequence as soon as  $\epsilon \in \text{Seq}$  is large enough and  $\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v$  has finite dimensional approximants (f.d.a) which is equivalent by [V1, Prop 3.3] to the stated f.d.a assumption (and then from the independence of the sequence of the value of the right hand side, for any sequence). We thus aim at proving the converse.

Fix  $n \in \mathbb{N}$  and then  $\delta > 0, K \in \mathbb{N}$  with

$$\liminf_{N \rightarrow \infty} \frac{1}{N^2} \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta, K, N; \epsilon) \leq \frac{1}{n} + \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \epsilon).$$

Define  $(\Xi_N, \Upsilon_N) \in \Gamma_{\infty, \emptyset}^U(\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v; \epsilon_N, \lfloor \epsilon_N^{-1} \rfloor, N)$  (assuming  $\epsilon$  large enough and the f.d.a. assumption) such that

$$\frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m}(\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N)) \right) \leq \frac{1}{n} + \frac{1}{N^2} \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_n | v; \delta, K, N).$$

First note that from the choice of  $\epsilon_N$  on deduces that

$$d(\tau_{(\Xi_N, \Upsilon_N)}, \tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v) \rightarrow_{N \rightarrow \infty} 0.$$

Now combining our previous inequalities and taking the  $\liminf_{N \rightarrow \infty}$ , one gets:

$$\begin{aligned} \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) &= \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \leq \\ \liminf_{N \rightarrow \infty} \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m}(\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N)) \right) &\leq \frac{2}{n} + \underline{\chi}_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \epsilon). \end{aligned}$$

And since the left hand side does not depend on  $n, \epsilon$  one can take an infimum in those variables to conclude. We will see soon that our orbital entropy depends only on  $W^*(\mathbf{U}_i)$  and the same proof gives we can consider self-adjoints variables as in [HMu]. Then the definition with sequences gives the equality.  $\square$

We gather in the next result the main properties of our new orbital free entropy.

**Proposition 8.8.** *The orbital free entropy satisfies the following properties.*

(1) *(Negativity and Vanishing)*

$$\chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) \leq 0, \quad \chi_{\text{orb}}(\mathbf{U}) = 0,$$

for any single multivariable  $\mathbf{U} = \{U_1, \dots, U_n\}$  having finite-dimensional approximants in the sens of [V1, Def 3.1].

(2) (*Monotonicity*)

$$\chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) \leq \chi_{orb}(\mathbf{V}_1; \dots; \mathbf{V}_m | v')$$

if  $\mathbf{V}_i \subset \mathbf{U}_i$  for  $1 \leq i \leq m, v' \subset v$ .

(3) (*Subadditivity*)

$$\chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_n, \mathbf{U}_{n+1}, \dots; \mathbf{U}_m | v) \leq \chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_n | v) + \chi_{orb}(\mathbf{U}_{n+1}, \dots, \mathbf{U}_m | v)$$

(4) (*Semicontinuity*)

If  $\mathbf{U}_1^{(k)} \sqcup \dots \sqcup \mathbf{U}_m^{(k)} \sqcup v^{(k)} \longrightarrow \mathbf{U}_1 \sqcup \dots \sqcup \mathbf{U}_m \sqcup v$  in distribution as  $k \rightarrow \infty$ , then

$$\chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) \geq \limsup_{k \rightarrow \infty} \chi_{orb}(\mathbf{U}_1^{(k)}, \dots, \mathbf{U}_m^{(k)} | v^{(k)}).$$

(5) (*Dependence on algebras*).

If  $\mathbf{U}_1, \dots, \mathbf{U}_n, \mathbf{V}_1, \dots, \mathbf{V}_n$  are multi-variables such that  $\mathbf{V}_i \subset W^*(\mathbf{U}_i)$  for  $1 \leq i \leq n$ ,  $v' \subset W^*(v)$ , then

$$\chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) \leq \chi_{orb}(\mathbf{V}_1, \dots, \mathbf{V}_m | v').$$

In particular,  $\chi_{orb}(\mathbf{U}_1, \dots, \mathbf{U}_m | v) = \chi_{orb}(W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_m) | W^*(v))$  depends only upon  $W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_m), W^*(v)$ .

(6) (*Orbital Talagrand's inequality and Characterization of Freeness*)

For  $\tau = \tau_{\mathbf{U}_1, \dots, \mathbf{U}_m}$ , let  $\tau_{free} = \tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m}$  the free product of its marginals, then :

$$d_W(\tau, \tau_{free}) \leq 4\sqrt{-\nu \chi_{orb}(\mathbf{U}_1; \dots; \mathbf{U}_m)}$$

where  $d_W$  is the 2-Wasserstein distance of [BV]. As a consequence, if  $\tau$  has finite-dimensional approximants, and if  $\chi_{orb}(\mathbf{U}_1; \dots; \mathbf{U}_m) = 0$  then  $\tau = \tau_{free}$ .

The converse of the last statement is contained in Theorem E and will be proved later.

*Proof.* (1) – (4) are straightforward.

(6) is valid for  $\bar{\chi}_{orb}$  with the proof of [Ue14, Appendix].

(5) We detail the proof of this statement since this property is pointed out in Theorem E, even though it is by now standard. Without loss of generality we assume that  $\mathbf{U}_i, \mathbf{V}_i, v = \mathbf{U}_{m+1}, v' = \mathbf{V}_{m+1}$  are all multivariables in  $\mathcal{U}(M)^\nu$  (This is possible since repeating variables clearly does not change orbital entropy).

Choose arbitrary  $K \in \mathbb{N}$  and  $\delta > 0$ . By the Kaplansky density theorem (in its unitary form [T, Th 4.11], jointly with the universal property of  $\mathcal{F}_1^\nu$ ), one can choose tuples of unitary continuous functions  $\mathbf{P}_i = (P_{ij})_{j=1}^\nu \in \prod_{j=1}^\nu \mathcal{F}_1^\nu, i = 1, \dots, n+1$  in such a way that  $\|V_{ij} - P_{ij}(\mathbf{U}_i)\|_2 < \delta/(4K2^{K-1})$ .

Then since non-commutative polynomials in  $\mathcal{F}_1^\nu$  are norm dense there, one can find  $\mathbf{Q}_i = (Q_{ij})_{j=1}^\nu \in \prod_{j=1}^\nu \mathbb{C}\langle \mathbf{U}_i \rangle$  such that  $\|Q_{ij}(u_2^1, \dots, u_2^\nu)\|_\infty \leq 2$  and  $\|Q_{ij}(u_2^1, \dots, u_2^\nu) - P_{ij}(u_2^1, \dots, u_2^\nu)\|_\infty < \delta/(4K2^{K-1})$ . This implies the corresponding inequality for free group variables  $u_2^1, \dots, u_2^\nu$  replaced by any family of unitaries, not only  $\mathbf{U}_i$  but also their matricial approximation.

Looking at the elements  $Q_{ij}(\mathbf{U}_i)$  (its degree and the sum of the absolute value of its coefficients only), one can also choose  $K' \in \mathbb{N}$  and  $\delta' > 0$  in such a way that if  $(\mathbf{A}_i)_{i=1}^{m+1} \in \Gamma_{\infty, \emptyset}^U(\mathbf{U}_1 \sqcup \dots \sqcup \mathbf{U}_m \sqcup v; \delta', K', N)$ , then

$$\left| \frac{1}{N} \text{Tr}(Q_{i_1 j_1}(\mathbf{A}_{i_1}) \cdots Q_{i_l j_l}(\mathbf{A}_{i_l})) - \tau(Q_{i_1 j_1}(\mathbf{U}_{i_1}) \cdots Q_{i_l j_l}(\mathbf{U}_{i_l})) \right| < \frac{\delta}{4}$$

for every  $1 \leq j_k \leq \nu, 1 \leq i_k \leq m, 1 \leq k \leq l$  and  $1 \leq l \leq K$ . The same  $K', \delta'$  reaches the same goal with  $\mathbf{U}_1 \sqcup \dots \sqcup \mathbf{U}_m \sqcup v$  replaced by the free product law. Remark here also that  $K', \delta'$  are independent of  $N$  by the way of finding them. If  $(\mathbf{A}_i)_{i=1}^{m+1} \in \Gamma_{\infty, \emptyset}^U(\mathbf{U}_1 \sqcup \dots \sqcup \mathbf{U}_m \sqcup v; \delta', K', N)$ ,

then

$$\begin{aligned}
& \left| \frac{1}{N} \text{Tr}(P_{i_1 j_1}(\mathbf{A}_{i_1}) \cdots P_{i_l j_l}(\mathbf{A}_{i_l})) - \tau(V_{i_1 j_1} \cdots V_{i_l j_l}) \right| \\
& \leq \frac{\delta}{4} + |\tau(P_{i_1 j_1}(\mathbf{U}_{i_1}) \cdots P_{i_l j_l}(\mathbf{U}_{i_l})) - \tau(V_{i_1 j_1} \cdots V_{i_l j_l})| \\
& + \left| \frac{1}{N} \text{Tr}(Q_{i_1 j_1}(\mathbf{A}_{i_1}) \cdots Q_{i_l j_l}(\mathbf{A}_{i_l})) - \frac{1}{N} \text{Tr}(P_{i_1 j_1}(\mathbf{A}_{i_1}) \cdots P_{i_l j_l}(\mathbf{A}_{i_l})) \right| \\
& + |\tau(Q_{i_1 j_1}(\mathbf{U}_{i_1}) \cdots Q_{i_l j_l}(\mathbf{U}_{i_l})) - \tau(P_{i_1 j_1}(\mathbf{U}_{i_1}) \cdots P_{i_l j_l}(\mathbf{U}_{i_l}))| \\
& \leq \frac{\delta}{4} + 3l2^{l-1} \max\{\|P_{ij}(\mathbf{X}_i) - Y_{ij}\|_2, \|Q_{ij} - P_{ij}\|_{\mathcal{F}_1^v} \mid 1 \leq j \leq \nu, 1 \leq i \leq m+1\} < \frac{\delta}{4} + 3\frac{\delta}{4} = \delta,
\end{aligned}$$

implying  $(\mathbf{P}_i(\mathbf{A}_i))_{i=1}^{m+1} = ((P_{ij}(\mathbf{A}_i))_{j=1}^{\nu})_{i=1}^{m+1} \in \Gamma_{\infty}(\mathbf{V}_1 \sqcup \cdots \sqcup \mathbf{V}_m \sqcup v'; \delta, K, N)$ . Hence, if  $(U_i)_{i=1}^n \in \Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : (\mathbf{A}_i)_{i=1}^m | \mathbf{A}_{m+1}; \delta', K', N)$ , i.e., (for  $U_{m+1} = I_N$ ),  $(U_i \mathbf{A}_i U_i^*)_{i=1}^{m+1} \in \Gamma_{\infty, \emptyset}^U(\mathbf{U}_1 \sqcup \cdots \sqcup \mathbf{U}_m \sqcup v; \delta', K', N)$ , then  $(U_i \mathbf{P}_i(\mathbf{A}_i) U_i^*)_{i=1}^{m+1} = (\mathbf{P}_i(U_i \mathbf{A}_i U_i^*))_{i=1}^{m+1} \in \Gamma_{\infty}(\mathbf{Y}_1 \sqcup \cdots \sqcup \mathbf{Y}_n; \delta, K, N)$  by the above consideration so that

$$(U_i)_{i=1}^m \in \Gamma_{\text{orb}}(\mathbf{V}_1, \dots, \mathbf{V}_m \sqcup v' : (\mathbf{P}_i(\mathbf{A}_i))_{i=1}^m | \mathbf{P}_{m+1}(\mathbf{A}_{m+1}); \delta, K, N).$$

Consequently, since the same kind of relation holds for  $(\mathbf{A}_i)_{i=1}^{m+1}$  in such a way that if they approximate the free product law,  $(\mathbf{P}_i(\mathbf{A}_i))_{i=1}^{m+1}$  approximate the corresponding free product law, one deduces:

$$\chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta', K', N) \leq \chi_{\text{orb}}(\mathbf{V}_1, \dots, \mathbf{V}_m | v'; \delta, K, N).$$

Hence the desired inequality immediately follows by the usual limits. The assertion on algebra dependence is immediate.  $\square$

**8.3. Additivity of orbital entropy: Proof of Theorem E.** The dependence on algebra in Theorem E has been obtained in proposition 8.8. We can and do assume that  $\mathfrak{U}_1, \dots, \mathfrak{U}_m, v$  have separately f.d.a (and thus so does their free product). Otherwise, both sides of equalities are clearly  $-\infty$ .

We call  $(\mathfrak{U}_1, \dots, \mathfrak{U}_m, v)$  variables with the free product trace:

$$\tau_{\mathfrak{U}_1, \dots, \mathfrak{U}_m, v} = \tau_{\mathbf{U}_1, \dots, \mathbf{U}_{m_1}} * \cdots * \tau_{\mathbf{U}_{m_n+1}, \dots, \mathbf{U}_m} * \tau_v.$$

We also set  $\widehat{\mathbf{U}}_1 = (\mathbf{U}_1, \dots, \mathbf{U}_{m_1}), \dots, \widehat{\mathbf{U}}_{n+1} = (\mathbf{U}_{m_n+1}, \dots, \mathbf{U}_m)$ .

We start by proving (1.2) and, first, part of the lower bound. This is the completely new argument that is enabled by Theorem 8.7.

### Step 1 : Lower bound using free product trace.

We want to prove for an appropriate non-increasing sequence  $\epsilon$  tending to 0:

$$\begin{aligned}
& \chi_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \\
(8.10) \quad & + \chi_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} | v, \epsilon) \leq \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}).
\end{aligned}$$

Take  $(\Xi_N, \Upsilon_N)$  as in the definition of  $\chi_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}})$ . Take any decreasing  $\eta_n \rightarrow 0$  and obtain:

$$\begin{aligned}
& \chi_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) \\
& \leq \liminf_{N \rightarrow \infty} \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} \left( \Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \eta_n, \lfloor \eta_n^{-1} \rfloor, N) \right) \right).
\end{aligned}$$

Then define by induction  $N_{n-1} < N_n$  such that

$$\begin{aligned}
& \chi_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) - \frac{1}{n} \\
& \leq \inf_{N \geq N_n} \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} \left( \Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \eta_n, \lfloor \eta_n^{-1} \rfloor, N) \right) \right).
\end{aligned}$$

Finally define  $n_N = n$  and  $\epsilon_N = \eta_n$  if  $N \in \llbracket N_n, N_{n+1} \rrbracket$  so that  $n_N \rightarrow \infty$ ,  $\epsilon_N \rightarrow 0$  (and is non-increasing) and for all  $N$ :

$$\begin{aligned} \chi_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) &- \frac{1}{n_N} \\ &\leq \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} (\Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N)) \right) \\ &\leq \frac{1}{N^2} \log \left( \gamma_{\mathcal{U}(M_N(\mathbb{C}))}^{\otimes m} (U\Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N)) \right), \end{aligned}$$

where we defined a unitary invariant version:

$$\begin{aligned} U\Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N) \\ := \{ (V_1 U_1, \dots, V_1 U_{m_1}, V_2 U_{m_1+1}, \dots, V_2 U_{m_2}, \dots, V_{n+1} U_m), V_i \in \mathcal{U}(M_N(\mathbb{C})) \\ (U_1, \dots, U_m) \in \Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N) \}. \end{aligned}$$

We fix  $\delta > 0, K \in \mathbb{N}$ , write for short  $U\Gamma_N := U\Gamma_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m : \Xi_N | \Upsilon_N; \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N)$ , and  $\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((\mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N) = \Gamma_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} : \Phi_N((\mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N)$ . We also set  $\gamma_{\mathcal{U}} = \gamma_{\mathcal{U}(M_N(\mathbb{C}))}$ . Define the probability measure  $\mathcal{P}$  on  $\mathcal{U}(M_N(\mathbb{C}))^m$  by:

$$\begin{aligned} \mathcal{P}(B) &= \int_{U\Gamma_N} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes(m-m_n)}(\mathbf{V}_{n+1}) \frac{1}{\gamma_{\mathcal{U}}^{\otimes n+1}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((\mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N))} \\ &\frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(U\Gamma_N)} \int_{\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((\mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N)} d\gamma_{\mathcal{U}}(U_1), \dots, d\gamma_{\mathcal{U}}(U_{n+1}) 1_B(U_1 \mathbf{V}_1, \dots, U_{n+1} \mathbf{V}_{n+1}) \\ &= \int_{(V_{i,1}^{-1} \mathbf{V}_i)_{i=1}^{n+1} \in U\Gamma_N} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes(m-m_n)}(\mathbf{V}_{n+1}) \frac{1}{\gamma_{\mathcal{U}}^{\otimes n+1}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((V_{i,1}^{-1} \mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N))} \\ &\frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(U\Gamma_N)} 1_{\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((V_{i,1}^{-1} \mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N)}(V_{i,1}, \dots, V_{(n+1),1}) 1_B(\mathbf{V}_1, \dots, \mathbf{V}_{n+1}). \end{aligned}$$

The last equality uses only change of variables and unitary invariance of Haar measure. Note that  $\mathcal{P}$  is supported in  $\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N)$  thus from the maximization for relative entropy recalled in section 2.8, one deduces :

$$\text{Ent}(\mathcal{P} | \gamma_{\mathcal{U}}^{\otimes m}) \leq \log \left( \gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N)) \right).$$

We now give a lower bound on this relative entropy. Note that the unitary invariant version  $U\mathcal{P}$  of  $\mathcal{P}$  by the action of  $\mathcal{U}(M_N(\mathbb{C}))^{n+1}$  given by  $(U_1, \dots, U_{n+1}) \mapsto (U_1 \mathbf{V}_1, \dots, U_{n+1} \mathbf{V}_{n+1})$  is nothing but  $U\mathcal{P} = \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(U\Gamma_N)} \gamma_{\mathcal{U}}^{\otimes m}|_{U\Gamma_N}$  by construction. Then a standard entropy computation (see e.g. [BD13, lemma 8.1] for a similar computation with a different action using the invariance of  $\gamma_{\mathcal{U}}^{\otimes m}$  for this action) gives:

$$\text{Ent}(\mathcal{P} | \gamma_{\mathcal{U}}^{\otimes m}) = \text{Ent}(\mathcal{P} | U\mathcal{P}) + \text{Ent}(U\mathcal{P} | \gamma_{\mathcal{U}}^{\otimes m}).$$

Of course, we have:  $\text{Ent}(U\mathcal{P} | \gamma_{\mathcal{U}}^{\otimes m}) = \log \left( \gamma_{\mathcal{U}}^{\otimes m}(U\Gamma_N) \right)$  and from the second formula for  $\mathcal{P}$  above computing the density with respect to  $U\mathcal{P}$ , one also deduces:

$$\begin{aligned} \text{Ent}(\mathcal{P} | U\mathcal{P}) &= \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(U\Gamma_N)} \int_{U\Gamma_N} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes(m-m_n)}(\mathbf{V}_{n+1}) \\ &\log \left( \gamma_{\mathcal{U}}^{\otimes n+1}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((V_{i,1}^{-1} \mathbf{V}_i)_{i=1}^{n+1}, \Xi_N) | \Upsilon; \delta, K, N)) \right) \\ &\geq \chi_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} | v; \delta, K, N; \epsilon) \end{aligned}$$

where we only used for the second line the unitary invariance of Haar measure to use the infimum definition uniformly over the probability measure on which we integrate. We thus proved:

$$\begin{aligned} & \log\left(\gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : \Xi_N | \Upsilon_N; \delta, K, N))\right) \\ & \geq \underline{\chi}_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} | v; \delta, K, N; \epsilon) + N^2 \underline{\chi}_{\text{orb}}(\mathfrak{U}_1, \dots, \mathfrak{U}_m | (\Xi_N)_{N \in \mathbb{N}}, (\Upsilon_N)_{N \in \mathbb{N}}) - \frac{N^2}{n_N}. \end{aligned}$$

Applying  $\lim_{K \rightarrow \infty} \liminf_{\delta \searrow 0} \frac{1}{N^2}$  concludes.

## Step 2 : Upper bound of (1.2).

The first upper bound of this type for orbital entropy was obtained in [BD13, Th 9.2]. Our following argument is similar to our previous argument there, but the new more constraining definition requires the use of the asymptotic freeness result of [V1].

Thus, fix  $\theta \in ]0, 1[$ . Fix,  $\delta > 0, K \in \mathbb{N}$ , we will consider  $\delta_1 < \delta, K_1 > K$  as in [V1, lemma 4.1] (we will explain this choice in function of  $\delta, K$  bellow). We start from the data in (8.3), namely fix  $(\Upsilon_1, \dots, \Upsilon_m, \Upsilon) \in \Gamma_{\infty, \emptyset}^U(\tau_{\mathbf{U}_1} * \dots * \tau_{\mathbf{U}_m} * \tau_v; \delta_1, K_1, N)$  and let  $\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon} = \Gamma_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m : (\Upsilon_i)_{i=1}^m, \Upsilon; \delta_1, K_1, N)$ . We now consider the probability measure:

$$\mathcal{P} = \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon})} \gamma_{\mathcal{U}}^{\otimes m} |_{\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon}}$$

and want to describe  $U\mathcal{P}$  the unitarily invariant version for the same action as in step 1. Of course, its support is :

$$\begin{aligned} U\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon} := \{ & (V_1 U_1, \dots, V_1 U_{m_1}, V_2 U_{m_1+1}, \dots, V_2 U_{m_2}, \dots, V_{n+1} U_m), V_i \in \mathcal{U}(M_N(\mathbb{C})) \\ & (U_1, \dots, U_m) \in \Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon} \}. \end{aligned}$$

$U\mathcal{P}$  is described as follows by definition and then change of variable and unitary invariance of Haar measure:

$$\begin{aligned} U\mathcal{P}(B) &= \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon})} \int_{\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon}} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes (m-m_n)}(\mathbf{V}_{n+1}) \\ & \quad \int d\gamma_{\mathcal{U}}(U_1) \dots d\gamma_{\mathcal{U}}(U_{n+1}) 1_B(U_1 \mathbf{V}_1, \dots, U_{n+1} \mathbf{V}_{n+1}) \\ &= \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon})} \int_{U\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon}} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes (m-m_n)}(\mathbf{V}_{n+1}) \\ & \quad 1_B(\mathbf{V}_1, \dots, \mathbf{V}_{n+1}) \gamma_{\mathcal{U}}^{\otimes (n+1)}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} : \Phi_N((\mathbf{V}_i)_{i=1}^m, (\Upsilon_i)_{i=1}^m) | \Upsilon; \delta_1, K_1, N)). \end{aligned}$$

Then one uses the same relation as in step 1 for entropy additivity and maximization given a support to deduce:

$$\text{Ent}(\mathcal{P} | \gamma_{\mathcal{U}}^{\otimes m}) = \text{Ent}(\mathcal{P} | U\mathcal{P}) + \text{Ent}(U\mathcal{P} | \gamma_{\mathcal{U}}^{\otimes m}) \leq \text{Ent}(\mathcal{P} | U\mathcal{P}) + \log\left(\gamma_{\mathcal{U}}^{\otimes m}(U\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon})\right).$$

Let us call  $\widehat{\Upsilon}_i$  the same group of variables as for  $\mathbf{U}_j$  starting from  $\Upsilon_j$ .

Since  $U\mathcal{P}$  is unitarily invariant and so is its image measure by  $\Phi_N(., (\widehat{\Upsilon}, \Upsilon))$ , one can use [V1, Corollary 2.14] to see that  $U\mathcal{P}(A) \geq 1 - \theta$  if

$$A = \{ \{ \mathbf{V}_1 \}, \dots, \{ \mathbf{V}_{n+1} \} \mid \Phi_N(\mathbf{V}_1, \widehat{\Upsilon}_1), \dots, \Phi_N(\mathbf{V}_{n+1}, \widehat{\Upsilon}_{n+1}), \{ \Upsilon \} \text{ are } (K_1, \delta_1) - \text{free} \}$$

and  $N$  is large enough (using for instance the decomposition of unitaries in real and imaginary part to reduce to the self-adjoint case considered in the quoted result). Finally, note that  $\delta_1, K_1$  can be chosen (independently of  $N$ ) such that

$$A \cap U\Gamma_{N, (\Upsilon_i)_{i=1}^m, \Upsilon} \subset \Gamma_{\text{orb}}(\tau_{\widehat{\mathbf{U}}_1} * \dots * \tau_{\widehat{\mathbf{U}}_{n+1}} * \tau_v : (\Upsilon_i)_{i=1}^m | \Upsilon; \delta, K, N)).$$

Thus, computing the relative entropy, we can estimate (since the integrand is non-positive, and  $\delta_1 < \delta, K_1 > K$  and the function under the first log is unitarily invariant by invariance of Haar measure):

$$\begin{aligned}
\text{Ent}(\mathcal{P}|\mathcal{U}\mathcal{P}) &= \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{N,(\mathbf{r}_i)_{i=1}^m, \Upsilon})} \int_{\Gamma_{N,(\mathbf{r}_i)_{i=1}^m, \Upsilon}} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes(m-m_n)}(\mathbf{V}_{n+1}) \\
&\quad \log \left( \gamma_{\mathcal{U}}^{\otimes(n+1)}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} : \Phi_N((\mathbf{V}_i)_{i=1}^m, (\mathbf{r}_i)_{i=1}^m) | \Upsilon; \delta_1, K_1, N)) \right) \\
&\leq \frac{1}{\gamma_{\mathcal{U}}^{\otimes m}(\Gamma_{N,(\mathbf{r}_i)_{i=1}^m, \Upsilon})} \int_{\Gamma_{N,(\mathbf{r}_i)_{i=1}^m, \Upsilon}} d\gamma_{\mathcal{U}}^{\otimes m_1}(\mathbf{V}_1) \dots d\gamma_{\mathcal{U}}^{\otimes(m-m_n)}(\mathbf{V}_{n+1}) \\
&\quad \int d\gamma_{\mathcal{U}}(U_1) \dots d\gamma_{\mathcal{U}}(U_{n+1}) \log \left( \gamma_{\mathcal{U}}^{\otimes(n+1)}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} : \Phi_N((U_i \mathbf{V}_i)_{i=1}^m, (\mathbf{r}_i)_{i=1}^m) | \Upsilon; \delta, K, N)) \right) \\
&\leq \int_{A \cap U\Gamma_{N,(\mathbf{r}_i)_{i=1}^m, \Upsilon}} d\mathcal{U}\mathcal{P}(\mathbf{V}_1, \dots, \mathbf{V}_{n+1}) \log \left( \gamma_{\mathcal{U}}^{\otimes(n+1)}(\Gamma_{\text{orb}}(\widehat{\mathbf{U}} : \Phi_N((\mathbf{V}_i)_{i=1}^m, (\mathbf{r}_i)_{i=1}^m) | \Upsilon; \delta, K, N)) \right) \\
&\leq \mathcal{U}\mathcal{P}(A) \times \chi_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} | v; \delta, K, N)
\end{aligned}$$

Since  $U\Gamma_{N,(\mathbf{r}_i)_{i=1}^m, \Upsilon}$  is easily included in a product, one obtains:

$$\begin{aligned}
\chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_m | v; \delta_1, K_1, N) &\leq (1 - \theta) \chi_{\text{orb}}(\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_{n+1} | v; \delta, K, N) \\
&\quad + \chi_{\text{orb}}(\mathbf{U}_1, \dots, \mathbf{U}_{m_1}; \delta_1, K_1, N) + \dots + \chi_{\text{orb}}(\mathbf{U}_{m_{n+1}}, \dots, \mathbf{U}_m; \delta_1, K_1, N).
\end{aligned}$$

Applying  $\lim_{\theta \rightarrow 0} \lim_{K \rightarrow \infty} \lim_{\delta_1 \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^2}$  concludes.

### Step 3 : Upper bound of (1.5).

This step is quite similar to step 2. Based on (2.16), (2.17), it suffices to show:

$$\tilde{\chi}_{\infty}(\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m) | v) \leq \chi_{\text{orb}}(W^*(\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m)), W^*(v)) + \tilde{\chi}_{\infty}(\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m)).$$

Thus, fix  $\theta \in ]0, 1[$ . Fix,  $\delta > 0, K \in \mathbb{N}$ , we will consider  $\delta_1 < \delta, K_1 > K$  as in [V1, lemma 4.1] (we will explain this choice in function of  $\delta, K$  bellow). We start from the data in Definition 2.8, namely fix  $\Upsilon \in \mathcal{U}(M_N(\mathbb{C}))$  with  $\tau_{\Upsilon} \in V_{\delta, K}(\tau_v)$  and let  $\Gamma_{N, \Upsilon} = \Gamma_{\infty, \Upsilon}^U(\tau_{\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m), v}, \delta_1, K_1, N)$ . We now consider the probability measure:  $\mathfrak{P} = \frac{1}{\Psi_* P(\Gamma_{N, \Upsilon})} \Psi_* P|_{\Gamma_{N, \Upsilon}}$  and want to describe  $\mathfrak{U}\mathfrak{P}$  the unitarily invariant version for the conjugation action of all the variables (note that  $\mathfrak{P}$  is not unitarily invariant because of the  $\Upsilon$ ). Of course, its support is  $\mathfrak{U}\Gamma_{N, \Upsilon}$  where for a set  $A$ :

$$\mathfrak{U}A := \{(VU_1V^*, \dots, VU_mV^*), V \in \mathcal{U}(M_N(\mathbb{C})) : (U_1, \dots, U_m) \in A\}.$$

$\mathfrak{U}\mathfrak{P}$  is described as follows by definition and then change of variable and unitary invariance of the gaussian measure  $P$  thus of  $\Psi_* P$  :

$$\begin{aligned}
\mathfrak{U}\mathfrak{P}(B) &= \frac{1}{\Psi_* P(\Gamma_{N, \Upsilon})} \int_{\Gamma_{N, \Upsilon}} d\Psi_* P(\mathbf{U}) \int d\gamma_{\mathcal{U}}(V) 1_B(V\mathbf{U}V^*) \\
&= \frac{1}{\Psi_* P(\Gamma_{N, \Upsilon})} \int_{\mathfrak{U}\Gamma_{N, \Upsilon}} d\Psi_* P(\mathbf{U}) 1_B(\mathbf{U}) \gamma_{\mathcal{U}}(\Gamma_{\text{orb}}(\tau_{\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m), v} : \mathbf{U} | \Upsilon; \delta_1, K_1, N)).
\end{aligned}$$

Then one uses the same relation as in step 1 for entropy additivity and maximization given a support to deduce:

$$\text{Ent}(\mathfrak{P}|\Psi_* P) = \text{Ent}(\mathfrak{P}|\mathfrak{U}\mathfrak{P}) + \text{Ent}(\mathfrak{U}\mathfrak{P}|\Psi_* P) \leq \text{Ent}(\mathfrak{P}|\mathfrak{U}\mathfrak{P}) + \log \left( \Psi_* P(\mathfrak{U}\Gamma_{N, \Upsilon}) \right).$$

Since  $\mathfrak{U}\mathfrak{P}$  and  $\mathfrak{U}\mathfrak{P} \times \delta_{\Upsilon}$  is unitarily invariant, one can use [V1, Corollary 2.14] to see that  $\mathfrak{U}\mathfrak{P}(\mathfrak{A}) \geq 1 - \theta$  if  $\mathfrak{A} = \{\mathbf{U} | \{\mathbf{U}\}, \{\Upsilon\} \text{ are } (K_1, \delta_1) - \text{free}\}$  and  $N$  is large enough. Finally, note that  $\delta_1, K_1$  can be chosen (independently of  $N$ ) such that  $\mathfrak{A} \cap \mathfrak{U}\Gamma_{N, \Upsilon} \subset \Gamma_{\infty, \Upsilon}^U(\tau_{\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m), v})^*$



$\tau_v|\Upsilon; \delta, K, N)$ ). Thus, computing the relative entropy, we can estimate as in step 2, in writing  $\Psi(\mathbf{X}) = (\Psi(\mathbf{X}_1), \dots, \Psi(\mathbf{X}_m))$ :

$$\text{Ent}(\mathfrak{P}|\mathfrak{U}\mathfrak{P}) \leq \mathfrak{U}\mathfrak{P}(A) \times \chi_{\text{orb}}(\Psi(\mathbf{X})|v; \delta, K, N)$$

Combining our inequalities, one obtains: and applying  $\lim_{\theta \rightarrow 0} \lim_{\delta \rightarrow 0} \lim_{K_1 \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N^2}$  concludes.

**Step 4 : Lower bound of (1.5).**

This step is similar to step 1. Take  $\Upsilon_N \in \mathcal{U}(M_N(\mathbb{C}))$  a sequence approximating  $v$  in law. Let  $\mathfrak{X}, v$  free variables with  $\mathfrak{X}$  having the same law as  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_m)$ .

We want to prove, for an appropriate non-increasing sequence  $\epsilon$  tending to 0:

$$(8.11) \quad \tilde{\chi}_{\infty}(\Psi(\mathfrak{X}) | (\Upsilon_N)_{N \in \mathbb{N}}) + \underline{\chi}_{\text{orb}}(\Psi(\mathbf{X})|v, \epsilon) \leq \tilde{\chi}_{\infty}(\Psi(\mathbf{X}) | (\Upsilon_N)_{N \in \mathbb{N}}).$$

Take any decreasing  $\eta_n \rightarrow 0$  and note that:

$$\tilde{\chi}_{\infty}(\Psi(\mathfrak{X}) | (\Upsilon_N)_{N \in \mathbb{N}}) \leq \liminf_{N \rightarrow \infty} \frac{1}{N^2} \log \left( \Psi_* P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \eta_n, \lfloor \eta_n^{-1} \rfloor, N)) \right).$$

Then define by induction  $N_{n-1} < N_n$  such that

$$\tilde{\chi}_{\infty}(\Psi(\mathfrak{X}) | (\Upsilon_N)_{N \in \mathbb{N}}) - \frac{1}{n} \leq \inf_{N \geq N_n} \frac{1}{N^2} \log \left( \Psi_* P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \eta_n, \lfloor \eta_n^{-1} \rfloor, N)) \right).$$

Finally define  $n_N = n$  and  $\epsilon_N = \eta_n$  if  $N \in \llbracket N_n, N_{n+1} \rrbracket$  so that  $n_N \rightarrow \infty$ ,  $\epsilon_N \rightarrow 0$  (and is non-increasing) and for all  $N$ :

$$\tilde{\chi}_{\infty}(\Psi(\mathfrak{X}) | (\Upsilon_N)_{N \in \mathbb{N}}) - \frac{1}{n_N} \leq \frac{1}{N^2} \log \left( \Psi_* P(\mathfrak{U}\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N)) \right).$$

where we used a unitary invariant version defined in step 3.

We fix  $\delta > 0, K \in \mathbb{N}$ , write for short  $\mathfrak{U}\Gamma_N := \mathfrak{U}\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \epsilon_n, \lfloor \epsilon_n^{-1} \rfloor, N)$  and  $\Gamma_{\text{orb}, N}(\mathbf{U}) := \Gamma_{\text{orb}}(\Psi(\mathbf{X}) : \mathbf{U} | \Upsilon_N; \delta, K, N)$ . Define the probability measure  $\mathbb{P}$  on  $\mathcal{U}(M_N(\mathbb{C}))^m$  by:

$$\begin{aligned} \mathbb{P}(B) &= \frac{1}{\Psi_* P(\mathfrak{U}\Gamma_N)} \int_{\mathfrak{U}\Gamma_N} d\Psi_* P(\mathbf{U}) \frac{1}{\gamma_{\mathcal{U}}(\Gamma_{\text{orb}, N}(\mathbf{U}))} \int_{\Gamma_{\text{orb}, N}(\mathbf{U})} d\gamma_{\mathcal{U}}(V) 1_B(VUV^*) \\ &= \frac{1}{\Psi_* P(\mathfrak{U}\Gamma_N)} \int_{\mathfrak{U}\Gamma_N} d\Psi_* P(\mathbf{U}) \frac{1}{\gamma_{\mathcal{U}}(\Gamma_{\text{orb}, N}(\mathbf{U}))} 1_B(\mathbf{U}) 1_{\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \delta, K, N)}(\mathbf{U}). \end{aligned}$$

Note that  $\mathbb{P}$  is supported in  $\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \delta, K, N)$  thus from the maximization for relative entropy recalled in section 2.8, one deduces :  $\text{Ent}(\mathbb{P} | \Psi_* P) \leq \log \left( \Psi_* P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \delta, K, N)) \right)$ .

We now give a lower bound on this relative entropy. Note that the unitary invariant version  $\mathfrak{U}\mathbb{P}$  of  $\mathbb{P}$  by the action of step 3 is nothing but  $\mathfrak{U}\mathbb{P} = \frac{1}{\Psi_* P(\mathfrak{U}\Gamma_N)} (\Psi_* P)|_{\mathfrak{U}\Gamma_N}$  by construction. Then the same computation as in step 1,3 gives:  $\text{Ent}(\mathbb{P} | \Psi_* P) = \text{Ent}(\mathbb{P} | \mathfrak{U}\mathbb{P}) + \text{Ent}(\mathfrak{U}\mathbb{P} | \Psi_* P)$ .

Of course, we have:  $\text{Ent}(\mathfrak{U}\mathbb{P} | \Psi_* P) = \log \left( \Psi_* P(\mathfrak{U}\Gamma_N) \right)$  and from the second formula for  $\mathbb{P}$  above computing the density with respect to  $\mathfrak{U}\mathbb{P}$ , one also deduces:

$$\text{Ent}(\mathbb{P} | \mathfrak{U}\mathbb{P}) = \frac{1}{\Psi_* P(\mathfrak{U}\Gamma_N)} \int_{\mathfrak{U}\Gamma_N} d\Psi_* P(\mathbf{U}) \log \left( \gamma_{\mathcal{U}}(\Gamma_{\text{orb}, N}(\mathbf{U})) \right) \geq \underline{\chi}_{\text{orb}}(\Psi(\mathbf{X})|v; \delta, K, N; \epsilon)$$

where we only used for the second line the unitary invariance of Haar measure to use the infimum definition uniformly over the probability measure on which we integrate. We thus proved:

$$\log \left( \Psi_* P(\Gamma_{\infty, \Upsilon_N}^U(\tau_{\mathfrak{X}, v}, \delta, K, N)) \right) \geq \underline{\chi}_{\text{orb}}(\Psi(\mathbf{X})|v; \delta, K, N; \epsilon) + N^2 \tilde{\chi}_{\infty}(\Psi(\mathfrak{X}) | (\Upsilon_N)_{N \in \mathbb{N}}) - \frac{N^2}{n_N}.$$

Applying  $\lim_{\delta \searrow 0} \liminf_{N \rightarrow \infty} \frac{1}{N^2}$  concludes.

**Step 5 : Equivalence of freeness and additivity and final bounds.**

Assuming as in (2) that the variables satisfy the additivity condition (1.3), then we have the vanishing orbital entropy (from the inequality in step 2) and thus freeness by Proposition 8.8 (6). Conversely, one shows as in [Ue14, Proposition 3.2] (only pay attention to gather the chosen matricial variables almost reaching the supremum in the definition of  $\chi_{\text{orb}}$  and gather them in an almost free bunch of matrices which is always possible with a supplementary use of [V1]), that if the stated variables are free then

$$\begin{aligned} \chi_{\text{orb}}(W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_m)) \\ \geq \underline{\chi}_{\text{orb}}(W^*(\mathbf{U}_1), \dots, W^*(\mathbf{U}_{m_1})) + \dots + \underline{\chi}_{\text{orb}}(W^*(\mathbf{U}_{m_n+1}), \dots, W^*(\mathbf{U}_m)) \end{aligned}$$

and from the equality  $\chi_{\text{orb}} = \underline{\chi}_{\text{orb}}$  in Theorem 8.7, one concludes to the additivity and thus to the vanishing orbital entropy condition in using again step 2.

The missing inequality to get (1.2) then follows from (8.10), Theorem 8.7 and the additivity in the free case applied on the first term. Similarly, the end of the proof of the lower bound from (1.5) and Theorems 8.7 and 7.2, missing in step 4, is the remark that  $\tilde{\chi}_{\infty}(\Psi(\mathfrak{X}) | (\Upsilon_N)_{N \in \mathbb{N}}) = \tilde{\chi}_{\infty}(\Psi(\mathbf{X}))$ . This is for instance a consequence of the proof of Proposition 2.18 with  $\{X_1 = Y\}, \{v\}, \{X_2\}$  free and  $X_2$  replaced by several variables  $\mathfrak{X}$  thanks to Theorem 7.2 so that one gets the equality

$$\chi(\mathfrak{X}, Y|v) = \chi(\mathfrak{X}|v) + \chi(Y) = \chi(\mathfrak{X}) + \chi(Y|v) = \chi(\mathfrak{X}) + \chi(Y)$$

(with the last equality coming  $\chi(Y) = \chi(Y|v)$  from the original [S02, Th 2.19]) which implies  $\chi(\mathfrak{X}) = \chi(\mathfrak{X}|v)$  and thus equality.

Finally, in case  $v = \emptyset$ , one can use (1.1) from [HMU, Theorem 2.6] and conclude from (1.2) to (1.4) (note the cases where some entropy is infinite are also easy and included in their result and using the usual subadditivity of free entropy one can reduce to the case where all the terms are finite). The case with  $v$  (when all entropies are finite the infinite case being easy using the general inequalities) can then be deduced from that case, Theorem 8.7 and (1.5), (1.2) as follows:

$$\begin{aligned} \chi(\mathbf{X}_1, \dots, \mathbf{X}_m|v) &= \chi_{\text{orb}}(W^*(\mathbf{X}_1, \dots, \mathbf{X}_m), W^*(v)) + \chi(\mathbf{X}_1, \dots, \mathbf{X}_m) \\ &= \chi_{\text{orb}}(W^*(\mathbf{X}_1, \dots, \mathbf{X}_m)|W^*(v)) + \chi(\mathbf{X}_1) + \dots + \chi(\mathbf{X}_m) + \chi_{\text{orb}}(W^*(\mathbf{X}_1), \dots, W^*(\mathbf{X}_m)) \\ &= \chi_{\text{orb}}(W^*(\mathbf{X}_1), \dots, W^*(\mathbf{X}_m)|W^*(v)) + \chi(\mathbf{X}_1) + \dots + \chi(\mathbf{X}_m). \end{aligned}$$

Similarly, using the same relations, one deduces the last relation (1.6)

$$\begin{aligned} \chi(\mathbf{X}_1|W^*(\mathbf{X}_2)) &= \chi_{\text{orb}}(W^*(\mathbf{X}_1), W^*(\mathbf{X}_2)) + \chi(\mathbf{X}_1) \\ &= \chi(\mathbf{X}_1, \mathbf{X}_2) - \chi(\mathbf{X}_2) - \chi(\mathbf{X}_1) + \chi(\mathbf{X}_1) \\ &= \chi(\mathbf{X}_1, \mathbf{X}_2) - \chi(\mathbf{X}_2). \end{aligned}$$

## REFERENCES

- [AGZ] G. ANDERSON. A. GUIONNET. and O. ZEITOUNI, *An introduction to random matrices* Cambridge University Press 2011.
- [BB] S.T. BELINSCHI and H. BERCOVICI, A property of free entropy, *Pacific J. Math.*, 211 (2003), 35–40.
- [BAG] G BEN AROUS AND A. GUIONNET Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy. *Probab. Theory Related Fields* 108, 4 (1997), 517–542.
- [BL] A. BENSOUSSAN and J-L. LIONS, *Applications des inéquations variationnelles en contrôle stochastique*. Dunod, Paris, 1978.
- [BCG] P. BIANE. A. GUIONNET. and M. CAPITAINÉ, Large deviation bounds for matrix brownian motion, *Invent. Math.* 152 (2003) 433–459.
- [BD13] P. BIANE. and Y. DABROWSKI, Concavification of free entropy. *Adv. Math.* 234 (2013) 667–696.
- [BS] P. BIANE and R. SPEICHER. Stochastic calculus with respect to free Brownian Motion. *Probability Theory and related Fields*, 112, no.3:373–409, 1998.
- [BV] P. BIANE, and D. VOICULESCU A free probability analogue of the Wasserstein metric on the trace-state space. *Geom. Funct. Anal.* 11 (2001), no. 6, 1125–1138.
- [BL] S.G. BOBKOV and M. LEDOUX. From Brunn–Minkowski To Brascamp–Lieb and to Logarithmic Sobolev Inequality *Geom. funct. anal.* Vol. 10 (2000) 1028–1052.
- [BD] M. BOUÉ and P. DUPUIS. A variational representation for certain functionals of Brownian motion, *Ann. Probab.* 26 (4) (1998) 1641–1659.

- [B] H. BRÉZIS. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert North-Holland (1973).
- [CDG1] T. CABANAL DUVILLARD, AND A. GUIONNET, Large deviations upper bounds for the laws of matrix-valued processes and non-communicative entropies. *Ann. Probab.* 29, 3 (2001), 1205–1261.
- [CDG2] CABANAL-DUVILLARD, T., AND GUIONNET, A. Discussions around Voiculescu’s free entropies. *Adv. Math.* 174, 2 (2003), 167–226.
- [CL] V. CAPRARO and M. LUPINI Introduction to Sofic and Hyperlinear Groups and Connes’ Embedding Conjecture LNM 2136 Springer, 2015.
- [Co76] A. CONNES: Classification of injective factors, *Ann. of Math.* (2) 104 (1976), 73–115.
- [D08] Y. DABROWSKI A note about proving non- $\Gamma$  under a finite non-microstate free Fisher information assumption. *JFA* 258 (2010), no 11, 3662–3674
- [D10] Y. DABROWSKI; A free stochastic partial differential equation, *Ann. Inst. Henri Poincaré Probab. Stat.*, 50 (2014), no. 4, 1404–1455.
- [D10b] Y. DABROWSKI A non-commutative Path Space approach to stationary free Stochastic Differential Equations, preprint arXiv:1006.4351.
- [D14] Y. DABROWSKI; Time Reversal of free diffusions I : Reversed Brownian motion, Reversed SDE and first order regularity of conjugate variables, arXiv:1402.4774
- [DDM] Y. DABROWSKI, K. DYKEMA and K. MUKHERJEE The simplex of tracial quantum symmetric states *Studia Mathematica* 225 (2014), 203–218.
- [DI] Y. DABROWSKI and A. IOANA Unbounded derivations, free dilations, and indecomposability results for  $II_1$  factors, *Trans. Amer. Math. Soc.*, 368 (2016) 4525–4560.
- [DZ] A. DEMBO and O. ZEITOUNI, *Large deviations techniques and applications*. Corrected reprint of the second (1998) edition. Stochastic Modelling and Applied Probability, 38, Springer-Verlag, Berlin, 2010.
- [DE] P. DUPUIS and R. ELLIS. *A weak convergence approach to the theory of large deviations* Wiley & Sons, New York (1997).
- [E] A. EBERHARD, Prox-regularity and subjets. In *Optimization and Related Topics, Applied Optimization*, Ch. 14, Vol. 47 (The Netherlands: Kluwer Academic Publishers), 2001 237–313.
- [ET] I. EKELAND and R. TÉMAM. *Convex analysis and variational problems* North-Holland (1976).
- [Fa] I. FARAH, B. HART and D. SHERMAN. Model theory of operator algebras II: Model theory. *Israel J. of Math* 201 (1), 477–505 (2014).
- [FU] D. FEYEL and A. S. ÜSTÜNEL The notion of convexity and concavity on Wiener space. *Journal of Functional Analysis*, 176, 400–428, 2000.
- [FS] W.H. FLEMING and H.M. SONER. Controlled Markov Processes and Viscosity Solutions Springer (2006).
- [Ga] M. GAO Free Markov processes and stochastic differential equations in von Neumann algebras. *Illinois Journal of Mathematics* 52, 1 (2008) 153–180.
- [Ge] I. GENTIL Dimensional contraction in Wasserstein distance for diffusion semigroups on a Riemannian manifold. *Potential Analysis* 42, 4 (2015), 861–873.
- [GM] A. GUIONNET and E. MAUREL-SÉGALA. Combinatorial aspects of matrix models, *ALEA Lat. Am. J. Probab. Math. Stat.* 1 (2006), 241–279.
- [GS09] A. GUIONNET and D. SHLYAKHTENKO, Free diffusions and matrix models with strictly convex interaction, *Geom. Funct. Anal.* 18 (2009), 1875–1916.
- [GS14] A. GUIONNET and D. SHLYAKHTENKO, Monotone free transport *Invent. Math.*, 197-3 (2014) 613–661.
- [GZ] A. GUIONNET and O. ZEITOUNI Large deviations asymptotics for spherical integrals. *J. Funct. Anal.* 188, 2 (2002), 461–515.
- [GZ2] A. GUIONNET and O. ZEITOUNI Addendum to: Large deviations asymptotics for spherical integrals. *J. Funct. Anal.* 216, 1 (2004), 230–241.
- [Ha] HARGÉ, G. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields* 130, 3 (2004), 415–440.
- [Hi] F. HIAI Free analog of pressure and its Legendre transform. *Comm. Math. Phys.* 255 (2005), no. 1, 229–252.
- [HMU] F. HIAI, T. MIYAMOTO and Y. UEDA, Orbital approach to microstates free entropy. *Int. J. Math.* 20 (2009), no. 2, 227–273.
- [HP] F. HIAI and D. PETZ. *The Semicircle Law, Free Random Variables and Entropy*, AMS, Mathematical surveys and monographs 77, 2000.
- [HP] J.B. HIRIART-URRUTY and P. PLAZANET. Moreau’s decomposition theorem revisited *Annales de l’I. H. P., section C*, tome S6 (1989), 325–338.
- [KS] I. KARATZAS and S.E. SHREVE *Brownian Motion and Stochastic Calculus*. Second Edition, Springer, New-York, 1998.
- [Ko] A.V. KOLESNIKOV On diffusion semigroups preserving the log-concavity. *J. Funct. Anal.*, 186 (2001), 196–205.
- [Kr] N. V. KRYLOV, On Kolmogorov’s equations for finite- dimensional diffusions, in *Stochastic PDE’s and Kolmogorov equations in infinite dimensions (Cetraro, 1998)*, *Lecture Notes in Math.*, vol. 1715, Springer, Berlin, 1999, 1–63.
- [KJF] A. KUFNER, O. JOHN and S. FUČÍK *Function spaces*. De Gruyter (1977).

- [K] S KULLBACK *Information Theory and Statistics*, Wiley publications in Statistics 1959.
- [L] J. LEHEC. Representation formula for the entropy and functional inequalities. *Ann. Inst. H. Poincaré Probab. Statist.* Volume 49, Number 3 (2013), 885–899.
- [LOS] J. LINDENSTRAUSS, G. OLSEN and Y. STERNFELD The Poulsen Simplex. *Ann. Inst. Fourier* 28 (1978), no. 1, 91–114.
- [LS] W. LIU, and M. STEPHAN Yosida approximations for multivalued stochastic partial differential equations driven by Lévy noise on a Gelfand triple. *J. Math. Anal. Appl.* 410 (2014) 158–178
- [M] G. W. MACKEY Induced Representations of Locally Compact Groups I *Annals of Mathematics*, Second Series, Vol. 55, No. 1 (Jan., 1952), 101–139
- [MS] I. MINEYEV and D. SHLYAKHTENKO. Non-microstate Free Entropy Dimension for Groups. *GAF*, 15:476–490, 2005.
- [MN] J. MINGO and A. NICA Annular non-crossing permutations and partitions, and second-order asymptotics for random matrices *Int. Math. Res. Not.*, 28 (2004), 1413–1460.
- [N75] S.M. NIKOL'SKII, *Approximation of functions of several variables and imbedding theorems*. Springer, (1975).
- [N06] D. NUALART, *The Malliavin Calculus and Related Topics*, Second Edition, Springer, (2006).
- [P] G. PISIER, *Introduction to Operator Space Theory*, Cambridge University Press, 2003.
- [PR] C. PRÉVOT and M. RÖCKNER *A concise course on stochastic partial differential equations* LNM 1905 Springer, 2007.
- [S] A. SOSHIKOV Universality at the edge of the spectrum in Wigner Random Matrices. *Com. Math. Phys.* 207 (1999) 697–733.
- [S02] D. SHLYAKHTENKO A Microstates Approach to relative free entropy *Int. J. Math.*, 13, 605 (2002) 605–623.
- [T] M. TAKESAKI, *Theory of operator algebras. I*. Reprint of the first (1979) edition. Encyclopaedia of Mathematical Sciences, 124. Operator Algebras and Non-commutative Geometry, 5. Springer-Verlag, Berlin, 2002.
- [Ue14] Y. UEDA, Orbital free entropy, revisited *Indiana Univ. Math. J.*, 63, 2 (2014), 551–577.
- [Us87] A. S. ÜSTÜNEL : Representation of the distributions on Wiener space and stochastic calculus of variations . *Journal of Functional Analysis*, 70 (1987), 126–139.
- [Us95] A. S. ÜSTÜNEL: *Introduction to Analysis on Wiener Space*. Lecture Notes in Math. Vol. 1610. Springer, 1995.
- [Us10] A. S. ÜSTÜNEL: *Analysis on Wiener Space and Applications*. <http://arxiv.org/abs/1003.1649>, 2010.
- [Us14] A. S. ÜSTÜNEL Variational calculation of Laplace transforms via entropy on Wiener space and applications *JFA* 267, 3058–3083 (2014).
- [V1] D. V. VOICULESCU A strengthened asymptotic freeness result for random matrices with applications to free entropy *International Mathematics Research Notices* 1998, No. 1, 41–63.
- [V2] D. V. VOICULESCU The analogues of entropy and of Fisher's information measure in free probability theory. II. *Invent. Math.* 118 (1994), no. 3, 411–440.
- [V3] D.V. VOICULESCU, The analogues of entropy and of Fisher's information measure in free probability theory, III : The absence of Cartan subalgebras, *Geometric and Functional Analysis* Vol 6, No. 1 (1996) 172–199.
- [V4] D. V. VOICULESCU The analogues of entropy and of Fisher's information measure in free probability theory. IV : Maximum entropy and freeness in Free Probability Theory, D.V. Voiculescu (ed.), Fields Inst. Commun. 12, Amer. Math. Soc., 1997, pp. 293–302.
- [V5] D.V. VOICULESCU. The analogues of entropy and of Fisher's information measure in free probability theory, V : non-commutative Hilbert Transforms. *Inven. Math.*, 132:189–227, 1998.
- [V6] D. V. VOICULESCU The analogues of entropy and of Fisher's information measure in free probability theory. VI. Liberation and Mutual Free Information *Adv. in Math.* 146 (1999), 101–166.
- [V] D.V. VOICULESCU, Free Entropy, *Bulletin of the London Mathematical Society* 34(3) (2002) 257–278.

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN UMR 5208, 43 BLVD. DU 11 NOVEMBRE 1918, F-69622 VILLEURBANNE CEDEX, FRANCE  
*E-mail address:* `dabrowski@math.univ-lyon1.fr`